

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	337	100.0	327	6	BD268215	Adenovirus
2	325	99.4	7469	6	BD268204	Adenovirus
3	325	99.4	7469	6	AX356037	Sequence
4	325	99.4	7469	6	BD021936	Adenovirus
5	325	99.4	10610	6	BD268212	Packaging
6	325	99.4	10610	6	AX356045	Sequence
7	325	99.4	10610	6	BD021944	Packaging
8	325	99.4	14455	6	BD268211	Adenovirus
9	325	99.4	14455	6	AX356044	Sequence
10	325	99.4	14455	6	BD021943	Packaging
11	309.2	94.6	5365	6	I09267	Sequence 34
12	309.2	94.6	5413	6	I09270	Sequence 37
13	309.2	94.6	5518	6	I09268	Sequence 35
14	309.2	94.6	5566	6	I09269	Sequence 36
15	309.2	94.6	6149	6	I09252	Sequence 19
16	309.2	94.6	6151	6	I09251	Sequence 18
17	281.4	86.1	4776	6	A95117	Sequence 1
18	281.4	86.1	4776	6	A95152	Sequence 1
19	281.4	86.1	4776	6	AR304364	Sequence 1

Query Match 100.0%; Score 327; DB 6; Length 327;  
Best Local Similarity 100.0%; Pred. No. 2.7e-66;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 60  
DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 60

QY 61 CCACTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGACCGAGGACC 120  
DB 61 CCACTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGACCGAGGACC 120

QY 121 TCAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACAGTCAC 180  
DB 121 TCAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACAGTCAC 180

QY 181 AGTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTCGGGTTGTTTC 240  
DB 181 AGTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTCGGGTTGTTTC 240

QY 241 TGGCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCG 300  
DB 241 TGGCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCG 300

QY 301 AGGTGAGGTGGCAGGCTTGAGATCT 327  
DB 301 AGGTGAGGTGGCAGGCTTGAGATCT 327

RESULT 2  
BD268204  
LOCUS Adenovirus vector, packaging cell line, composition and method for production and use. 7469 bp DNA linear PAT 17-JUL-2003  
DEFINITION BD268204  
ACCESSION BD268204.1 GI:33077972  
VERSION JP 2002534130-A/8.  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 7469)  
AUTHORS Nemerow, G.R.; Seggern, D.J.V., Hallenbeck, P.L., Stevenson, S.C. and Skripchenko, Y.  
TITLE Adenovirus vector, packaging cell line, composition and method for production and use  
JOURNAL Patent: JP 2002534130-A 8 15-OCT-2002;  
COMMENT NOVARTIS AG, THE SCRIPPS RESEARCH INSTITUTE  
OS Artificial Sequence  
PN JP 2002534130-A/8  
PD 15-OCT-2002  
PF 14-JAN-2000 JP 2000593765  
PR 14-JAN-1999 US 60/115920  
PI GLEN ROBERT NEMEROW, DANIEL J VON SEGGERN, PAUL L HALLENBECK, PI SUSAN C STEVENSON, YELENA SKRIPCHENKO  
PC C12N15/09, A61K35/76, A61K48/00, A61P35/00, A61P43/00, A61P43/00, C12N5/10,  
PC C12N7/00, C12Q1/68, G01N33/53, G01N33/566, C12N15/00, C12N5/00 CC  
Description of Artificial Sequence: plasmid  
FH key Location/Qualifiers  
FT source 1..7469  
FT /organism='Artificial Sequence'.  
FEATURES  
source  
1..7469  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 99.4%; Score 325; DB 6; Length 7469;  
Best Local Similarity 100.0%; Pred. No. 7.7e-66;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 61

DB 908 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 967

QY 62 CAGTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGCCACCGAGGACCT 121  
DB 968 CAGTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCACA 181  
DB 1028 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCACA 1087

QY 182 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTTGTTTCT 241  
DB 1088 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTTGTTTCT 1147

QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 301  
DB 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 1207

QY 302 GGTGAGGTGGCAGGCTTGAGATC 326  
DB 1208 GGTGAGGTGGCAGGCTTGAGATC 1232

RESULT 3  
AX356037  
LOCUS Sequence 8 from Patent WO0183729. 7469 bp DNA linear PAT 06-FEB-2002  
DEFINITION AX356037  
ACCESSION AX356037  
VERSION AX356037.1 GI:18620599  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Nemerow, G.R., von Seggern, D.J. and Friedlander, M.  
TITLE Vectors for ocular transduction and use thereof for genetic therapy  
JOURNAL Patent: WO 0183729-A 8 08-NOV-2001;  
COMMENT NOVARTIS AG (CH); The Scripps Research Institute (US); Nemerow, Glen R. (US); Von Seggern, Daniel J. (US); Friedlander, Marty (US)  
FEATURES  
source  
1..7469  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="plasmid"

ORIGIN  
Query Match 99.4%; Score 325; DB 6; Length 7469;  
Best Local Similarity 100.0%; Pred. No. 7.7e-66;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 61  
DB 908 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 967

QY 62 CAGTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGCCACCGAGGACCT 121  
DB 968 CAGTACTCTTTGGATCGGAAACCCGTCGCGCTCCGAAACCGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCACA 181  
DB 1028 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCACA 1087

QY 182 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTTGTTTCT 241  
DB 1088 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGCGGCTGGCGGTTGTTTCT 1147

QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 301  
DB 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 1207



QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326  
 |||||  
 Db 1208 GGTGAGGTGTGGCAGGCTTGAGATC 1232

RESULT 4  
 BD021936 7469 bp DNA linear PAT 27-AUG-2002  
 LOCUS  
 DEFINITION Packaging cell systems for use in promotion of the development of high-capacity adenoviral vectors.

ACCESSION BD021936  
 VERSION BD021936.1 GI:22563159  
 KEYWORDS JP 2001505047-A/8.  
 SOURCE unclassified  
 ORGANISM unclassified

REFERENCE 1 (bases 1 to 7469)  
 AUTHORS Memrow,G.R. and Seggern,D.J.V.  
 TITLE Packaging cell systems for use in promotion of the development of high-capacity adenoviral vectors  
 JOURNAL Patent: JP 2001505047-A 8 17-APR-2001;  
 NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE  
 COMMENT PN JP 2001505047-A/8  
 PD 17-APR-2001  
 PF 24-SEP-1997 JP 1998515273  
 PR 25-SEP-1996 US 08/719806  
 PI GLEN R MEMROW DANIEL J VON SEGGERN  
 PC C12N5/10,C07K14/075,C12N15/09/A61K31/711,A61K35/76,A61K48/00,  
 A61P35/00,  
 PC C12N5/00,C12N15/00  
 PC C12N5/00,C12N15/00  
 CC Strandedness: Double;  
 CC Topology: Circular;  
 FH Key Location/Qualifiers

FEATURES  
 source  
 1..7469  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 99.4%; Score 325; DB 6; Length 7469;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-66;  
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGCTTTTC 61  
 |||||  
 Db 908 GATCTGAATTCGAGCTCGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGCTTTTC 967

QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 121  
 |||||  
 Db 968 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCGTCTAACCACTCACA 181  
 |||||  
 Db 1028 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCGTCTAACCACTCACA 1087

QY 182 GTCCGAAGTAGCTGAGCACCGTCGGCGGCGGACCGGTGCGGCTCGGGTCTTTCT 241  
 |||||  
 Db 1088 GTCCGAAGTAGCTGAGCACCGTCGGCGGCGGACCGGTGCGGCTCGGGTCTTTCT 1147

QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGCTCGA 301  
 |||||  
 Db 1148 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGCTCGA 1207

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326  
 |||||  
 Db 1208 GGTGAGGTGTGGCAGGCTTGAGATC 1232

RESULT 5  
 BD268212 10610 bp DNA linear PAT 17-JUL-2003  
 LOCUS  
 DEFINITION Adenovirus vector, packaging cell line, composition and method for

production and use.  
 BD268212  
 VERSION BD268212.1 GI:33077980  
 KEYWORDS JP 2002534130-A/16.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 10610)  
 AUTHORS Memrow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.  
 TITLE Adenovirus vector, packaging cell line, composition and method for production and use  
 JOURNAL Patent: JP 2002534130-A 16 15-OCT-2002;  
 NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE  
 COMMENT OS Artificial Sequence  
 PN JP 2002534130-A/16  
 PD 15-OCT-2002  
 PF 14-JAN-2000 JP 2000593765  
 PR 14-JAN-1999 US 60/115920  
 PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCHENKO  
 PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,  
 C12N5/10,  
 PC C12N5/10,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC  
 PC C12N7/00,Description of Artificial Sequence: plasmid  
 FH Key Location/Qualifiers  
 FT source 1..10610  
 /organism="Artificial Sequence".  
 /organism="Artificial Sequence".  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

ORIGIN  
 Query Match 99.4%; Score 325; DB 6; Length 10610;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-66;  
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGCTTTTC 61  
 |||||  
 Db 4049 GATCTGAATTCGAGCTCGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGCTTTTC 4108

QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 121  
 |||||  
 Db 4109 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 4168

QY 122 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCGTCTAACCACTCACA 181  
 |||||  
 Db 4169 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCGTCTAACCACTCACA 4228

QY 182 GTCCGAAGTAGCTGAGCACCGTCGGCGGCGGACCGGTGCGGCTCGGGTCTTTCT 241  
 |||||  
 Db 4229 GTCCGAAGTAGCTGAGCACCGTCGGCGGCGGACCGGTGCGGCTCGGGTCTTTCT 4288

QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGCTCGA 301  
 |||||  
 Db 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGCTCGA 4348

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326  
 |||||  
 Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 6  
 AX356045 10610 bp DNA linear PAT 06-FEB-2002  
 LOCUS  
 DEFINITION Sequence 16 from Patent WO0183729.  
 ACCESSION AX356045  
 VERSION AX356045.1 GI:18620607  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM synthetic construct

```
other sequences; artificial sequences.
1
REFERENCE
  AUTHORS      Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  TITLE        Vectors for ocular transduction and use thereof for genetic therapy
  JOURNAL      Patent: WO 0183729-A 16 08-NOV-2001;
                Novartis AG (CH) ; The Scripps Research Institute (US) ; Nemerow,
                Glen R. (US) ; Von Seggern, Daniel J. (US) ; Friedlander, Marty
                (US)
FEATURES
  source       Location/Qualifiers
                1..10610
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="plasmid"
ORIGIN
  Query Match 99.4%; Score 325; DB 6; Length 10610;
  Best Local Similarity 100.0%; Pred. No. 7.6e-66;
  Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTC 61
  Db 4049 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTC 4108
  QY 62 CAGTACTCTTGATCGGAACCCGTCGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 121
  Db 4109 CAGTACTCTTGATCGGAACCCGTCGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 4168
  QY 122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAGAAAGCGGTCTAACCACTCACA 181
  Db 4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAGAAAGCGGTCTAACCACTCACA 4228
  QY 182 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGTGCGGTTGTTTCT 241
  Db 4229 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGTGCGGTTGTTTCT 4288
  QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGATGGTGA 301
  Db 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGATGGTGA 4348
  QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
  Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373
RESULT 7
BD021944 10610 bp DNA linear PAT 27-AUG-2002
LOCUS     Packaging cell systems for use in promotion of the development of
DEFINITION high-capacity adenoviral vectors.
ACCESSION BD021944
VERSION    BD021944.1 GI:22563167
KEYWORDS  JP 2001505047-A/16.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 10610)
AUTHORS    Nemerow,G.R. and Seggern,D.J.V.
TITLE      Packaging cell systems for use in promotion of the development of
JOURNAL    high-capacity adenoviral vectors
COMMENT    Patent: JP 2001505047-A 16 17-APR-2001;
            NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
            PN JP 2001505047-A/16
            PD 17-APR-2001
            PF 24-SEP-1997 JP 1998515273
            PR 25-SEP-1996 US 08/719806
            PI GLEN R. NEMEROW,DANIEL J VON SEGGERN
            PC C12N5/10,C07K14/075,C12N15/09//A61K31/711,A61K35/76,A61K48/00,
            PG A61P35/00,
            PC C12N5/00,C12N15/00
            CC Strandedness: Double;
            CC Topology: Circular;
            PH Key Location/Qualifiers.
            1..14455
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
other sequences; artificial sequences.
1
REFERENCE
  AUTHORS      Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  TITLE        Vectors for ocular transduction and use thereof for genetic therapy
  JOURNAL      Patent: WO 0183729-A 16 08-NOV-2001;
                Novartis AG (CH) ; The Scripps Research Institute (US) ; Nemerow,
                Glen R. (US) ; Von Seggern, Daniel J. (US) ; Friedlander, Marty
                (US)
FEATURES
  source       Location/Qualifiers
                1..10610
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
ORIGIN
  Query Match 99.4%; Score 325; DB 6; Length 10610;
  Best Local Similarity 100.0%; Pred. No. 7.6e-66;
  Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTC 61
  Db 4049 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTC 4108
  QY 62 CAGTACTCTTGATCGGAACCCGTCGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 121
  Db 4109 CAGTACTCTTGATCGGAACCCGTCGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 4168
  QY 122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAGAAAGCGGTCTAACCACTCACA 181
  Db 4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAGAAAGCGGTCTAACCACTCACA 4228
  QY 182 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGTGCGGTTGTTTCT 241
  Db 4229 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGTGCGGTTGTTTCT 4288
  QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGATGGTGA 301
  Db 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGATGGTGA 4348
  QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
  Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373
RESULT 8
BD268211/c 14455 bp DNA linear PAT 17-JUL-2003
LOCUS     Adenovirus vector, packaging cell line, composition and method for
DEFINITION production and use.
ACCESSION BD268211
VERSION    BD268211.1 GI:33077979
KEYWORDS  JP 2002534130-A/15.
SOURCE     synthetic construct
ORGANISM   other sequences; artificial sequences.
REFERENCE  1 (bases 1 to 14455)
AUTHORS    Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
            Skripchenko,Y.
TITLE      Adenovirus vector, packaging cell line, composition and method for
JOURNAL    production and use
COMMENT    Patent: JP 2002534130-A 15 15-OCT-2002;
            NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
            OS Artificial Sequence
            PN JP 2002534130-A/15
            PD 15-OCT-2002
            PF 14-JAN-1999 US 60/115920
            PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
            SUSAN C STEVENSON,YELENA SKRIPCHENKO
            PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,
            PC C12N5/10,
            PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence: plasmid
            FH Key Location/Qualifiers
            FT source 1..14455
            /organism='Artificial Sequence'.
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
```

```
ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 241
DB 13135 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 13076

QY 242 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 301
DB 13075 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 9
AX356044/c
LOCUS AX356044 14455 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 15 from Patent WO0183729.
ACCESSION AX356044
VERSION AX356044.1 GI:18620606
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
TITLES Vectors for ocular transduction and use thereof for genetic therapy
JOURNAL Patent: WO 0183729-A 15 08-NOV-2001;
Novartis AG (CH); The Scripps Research Institute (US); Nemerow,
Glen R. (US); Von Seggern, Daniel J. (US); Friedlander, Marty
(US)
FEATURES
source
Location/Qualifiers
1. .14455
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="plasmid"

ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 241
DB 13135 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 13076

QY 242 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 301
DB 13075 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 9
AX356044/c
LOCUS AX356044 14455 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 15 from Patent WO0183729.
ACCESSION AX356044
VERSION AX356044.1 GI:18620606
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Nemerow,G.R., von Seggern,D.J. and Seggern,D.J.V.
TITLES Packaging cell systems for use in promotion of the development of
high-capacity adenoviral vectors
JOURNAL Patent: JP 2001505047-A 15 17-APR-2001;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
PN JP 2001505047-A/15
PD 17-APR-2001
PF 24-SEP-1997 JP 1998515273
PI GLEN R NEMEROW DANIEL J VON SEGGERN
PC C12N5/10,C07K14/075,C12N15/09//A61K31/711,A61K35/76,A61K48/00,
PC A61P35/00,
PC C12N5/00,C12N15/00
CC Strandedness: Double;
CC Topology: Circular;
FH Key Location/Qualifiers
1. .14455
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTTCGGGTTGAGGACAACTCTTCGGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 241
DB 13135 GTCCAAAGGTAGGCTGAGCACCCTGCGGGGGGCGGAGCGGTCGGGGTTGTTTCT 13076

QY 242 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 301
DB 13075 GGCGGAGTCTGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGGTCA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991
```

```
Db 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 11
LOCUS I09267 5365 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 34 from Patent WO 8901940.
ACCESSION I09267
VERSION I09267.1 GI:588051
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5365)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 34 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5365
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5365;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTTGGGCTCGCGGTCGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924

RESULT 12
LOCUS I09270 5413 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 37 from Patent WO 8901940.
ACCESSION I09270
VERSION I09270.1 GI:588054
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5413)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 37 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5413
/organism="unknown"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5518;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924

RESULT 13
LOCUS I09268 5518 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 35 from Patent WO 8901940.
ACCESSION I09268
VERSION I09268.1 GI:588052
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5518)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 35 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5518
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5518;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTGTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATCGTTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 461.032 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-26  
Perfect score: 327  
Sequence: 1 agatctgaattcgagctgc.....gtgtggcaggcttgatctct 327

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	327	3	AAQ59054 Nucleotid
2	327	100.0	327	6	ABA94266 Adenoviru
3	327	100.0	327	10	ADB75112 Adenoviru
4	327	100.0	327	10	ADF48736 Adenoviru
5	325	99.4	7469	2	AAV32372 Complete
6	325	99.4	7469	3	AAQ59043 Nucleotid
7	325	99.4	7469	6	ABA94252 Nucleotid
8	325	99.4	7469	10	ADB75149 Adenoviru
9	325	99.4	7469	10	ADF48794 Fibre exp
10	325	99.4	10610	2	AAV32375 Complete
11	325	99.4	10610	3	AAQ59051 Nucleotid
12	325	99.4	10610	6	ABA94260 Nucleotid
13	325	99.4	10610	10	ADB75157 Adenoviru
14	325	99.4	10610	10	ADF48802 Complete
15	325	99.4	14455	2	AAV32374 Complete
16	325	99.4	14455	3	AAQ59050 Nucleotid
17	325	99.4	14455	6	ABA94259 Nucleotid
18	325	99.4	14455	10	ADB75156 Adenoviru
19	325	99.4	14455	10	ADF48801 E1/fibre
20	310.2	94.9	7316	2	AAQ04555 Plasmid p

21	310.2	94.9	7377	2	AAQ05607	Ad12250	CMV5 prom
22	310.2	94.9	7377	2	AAQ03005	Ad12252	CMV5 prom
23	309.2	94.6	348	13	ADR12252	Ad12254	CMV5 prom
24	309.2	94.6	531	13	ADR12254	Ad12255	CMV5 prom
25	309.2	94.6	3953	13	ADR12255	Ad12255	CMV5 prom
26	309.2	94.6	5413	1	AAQ90649	Ad12255	CMV5 prom
27	309.2	94.6	5518	1	AAQ90647	Ad12255	CMV5 prom
28	309.2	94.6	5566	1	AAQ90648	Ad12255	CMV5 prom
29	309.2	94.6	6051	2	AAQ14934	Ad12255	CMV5 prom
30	309.2	94.6	6149	1	AAQ90645	Ad12255	CMV5 prom
31	309.2	94.6	6151	1	AAQ90644	Ad12255	CMV5 prom
32	309.2	94.6	6151	2	AAQ05608	Ad12255	CMV5 prom
33	309.2	94.6	6151	2	AAQ03006	Ad12255	CMV5 prom
34	309.2	94.6	6165	2	AAQ20324	Ad12255	CMV5 prom
35	307.6	94.1	5365	1	AAQ90646	Ad12255	CMV5 prom
36	306	93.6	8073	3	AAA64553	Ad12255	CMV5 prom
37	296.2	90.6	335	13	ADR12250	Ad12255	CMV5 prom
38	281.4	86.1	4776	2	AAQ77617	Ad12255	CMV5 prom
39	281.4	86.1	4776	2	AAQ77614	Ad12255	CMV5 prom
40	271	82.9	1115	13	ADR12257	Ad12255	CMV5 prom
C 41	203	62.1	30365	6	ABK49011	Ad12255	CMV5 prom
C 42	203	62.1	31672	6	ABK49010	Ad12255	CMV5 prom
C 43	203	62.1	34616	6	ABK49009	Ad12255	CMV5 prom
44	203	62.1	35937	6	ABK69881	Ad12255	CMV5 prom
45	203	62.1	35937	9	ACC70007	Ad12255	CMV5 prom

ALIGNMENTS

RESULT 1  
AAA59054  
ID AAA59054 standard; DNA; 327 BP.  
XX  
AC AAA59054;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE Nucleotide sequence of a tripartite leader sequence.  
XX  
KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;  
KW ss.  
XX  
OS Mastadenovirus.  
XX  
PN WO200042208-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 14-JAN-2000; 2000WO-EP0000265.  
XX  
PR 14-JAN-1999; 99US-0115920P.  
XX  
PA (NOVS ) NOVARTIS AG.  
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
XX (SCRI ) SCRIPPS RES INST.  
PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;  
PI Skripchenko Y;  
XX  
WPI; 2000-476068/41.  
XX  
PT New nucleic acid comprising an adenovirus tripartite leader nucleotide  
PT for producing high-capacity and targeted vectors for adenovirus-based  
PT gene therapy.  
XX  
PS Claim 14; Page 169; 212pp; English.  
XX  
CC The specification describes a nucleic acid molecule comprising an  
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence  
CC comprising two different TPL exons or three same or different TPL exons.  
CC The nucleic acid is used to produce an adenovirus vector particle,  
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral

CC vectors, target an adenovirus vector to a cell, produce a modified  
CC adenovirus, deliver a heterologous gene to an animal and produce a  
CC gutless adenoviral vector particle. The present sequence represents a TPL  
CC sequence, which is used to construct nucleic acid molecules of the  
CC invention

SQ Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;  
Query Match 100.0%; Score 327; DB 3; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.8e-80;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60  
DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60  
QY 61 CCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACC 120  
DB 61 CCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACC 120  
QY 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180  
DB 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180  
QY 181 AGTCCCAAGTAGGCTGAGCACCGTGGCGGCGGAGCGGTGGCGGTTCGCGGTGTTTC 240  
DB 181 AGTCCCAAGTAGGCTGAGCACCGTGGCGGCGGAGCGGTGGCGGTTCGCGGTGTTTC 240  
QY 241 TGGCGGAGTGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTCG 300  
DB 241 TGGCGGAGTGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTCG 300  
QY 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327  
DB 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327

RESULT 2  
ABA94266  
ID ABA94266 standard; DNA; 327 BP.  
XX ABA94266;  
XX  
XX  
DT 07-AUG-2003 (revised)  
DT 13-MAR-2002 (first entry)  
XX  
DE Adenovirus 5 tripartite leader (TPL) partial nucleotide sequence.  
XX  
XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;  
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; tripartite leader; TPL; ds.  
XX  
OS Human adenovirus type 5.  
XX  
XX WO200183729-A2.  
XX  
XX  
XX 08-NOV-2001.  
XX  
XX 30-APR-2001; 2001WO-EP004863.  
XX  
XX 01-MAY-2000; 2000US-00562934.  
XX  
XX (NOVS ) NOVARTIS AG.  
PA (SCRI ) SCRIPPS RES INST.  
PA (NEME/) NEMEROW G R.  
PA (VSEGE/) VON SEGGERN D J.  
PA (FRIE/) FRIEDLANDER M.  
XX  
XX Nemerow GR, Von Seggern DJ, Friedlander M;  
XX  
XX WPI; 2002-082846/11.  
XX

PT Polynucleotide for making vectors, useful for treating ocular diseases,  
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat  
XX sequences, packaging signal and photoreceptor-specific promoter.  
XX  
PS Example 1; Page 122; 149pp; English.

XX  
CC The invention provides an isolated polynucleotide comprising adenovirus  
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal  
CC operatively linked to ITRS and a photoreceptor-specific promoter. A  
CC recombinant AV vector (AAV) comprising the polynucleotide is useful for  
CC targeted delivery of a gene product to the eye (especially to the  
CC vitreous cavity), for treating an ocular disease, e.g., retinal  
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic  
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal  
CC preferably human. The AAV comprises a fiber protein that specifically or  
CC selectively binds to receptors that are expressed on cells (preferably  
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a  
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein  
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5  
CC penton, and the therapeutic product is a trophic factor, an anti-  
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type  
CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that  
CC regulates expression of a photoreceptor specific gene product. The viral  
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV  
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful  
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber  
CC protein or its portion, and selectively transduces photoreceptors and  
CC delivers a gene product encoded by AAV. The present sequence represents a  
CC adenovirus 5 tripartite leader (TPL) partial nucleotide sequence. (Updated  
CC on 07-AUG-2003 to correct OS field.)  
XX

SQ Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;  
Query Match 100.0%; Score 327; DB 6; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.8e-80;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60  
DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60  
QY 61 CCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCGAACGGTACTCCGCCACCGAGGGACC 120  
DB 61 CCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCGAACGGTACTCCGCCACCGAGGGACC 120  
QY 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180  
DB 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180  
QY 181 AGTCCCAAGTAGGCTGAGCACCGTGGCGGCGGAGCGGTGGCGGTTCGCGGTGTTTC 240  
DB 181 AGTCCCAAGTAGGCTGAGCACCGTGGCGGCGGAGCGGTGGCGGTTCGCGGTGTTTC 240  
QY 241 TGGCGGAGTGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTCG 300  
DB 241 TGGCGGAGTGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTCG 300  
QY 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327  
DB 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327

RESULT 3  
ADB75112  
ID ADB75112 standard; DNA; 327 BP.  
XX  
XX  
AC ADB75112;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
XX Adenovirus type 5 tripartite leader sequence #1.  
XX  
XX ophthalmological; antiinflammatory; antidiabetic; gene therapy;  
KW



adenovirus inverted terminal repeat sequence;  
 adenovirus packaging signal; photoreceptor-specific promoter;  
 adenovirus type 37; adenovirus type D serotype; adenovirus type 2;  
 adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;  
 rhodopsin; wild-type Stargardt disease gene; STDG1; anti-cancer agent;  
 retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;  
 diabetic retinopathy; retinal vascularisation; choroideraemia;  
 gyrate atrophy; macular dystrophy; retinoblastoma;  
 photoreceptor-restricted transgene expression;  
 recombinant adenovirus vector; adenovirus type 5; Ad5;  
 tripartite leader sequence; TPL; ds.

Human adenovirus type 5.

US2002193327-A1.

19-DEC-2002.

01-MAY-2001; 2001US-00847101.

01-MAY-2000; 2000US-00562934.

(SCRI ) SCRIPPS RES INST.

Nemerow GR, Von Seggern DJ, Friedlander M;

WPI; 2003-657234/62.

Novel nucleic acids comprising adenovirus inverted terminal repeat sequences, adenovirus packaging signals operatively linked to the sequences and photoreceptor-specific promoters, useful for treating retinitis pigmentosa.

Example 1; Page 62; 106pp; English.

The invention describes an isolated nucleic acid (I) comprising adenovirus inverted terminal repeat sequence, an adenovirus packaging signal operatively linked to the sequence, and a photoreceptor-specific promoter. A Recombinant adenovirus vector (II) comprising (I) is useful for targeted delivery of a gene product to the eye of a mammal which involves administering (II) that comprises heterologous DNA encoding the gene product or resulting in expression of the gene product, where the recombinant virus comprises a fibre protein that specifically or selectively binds to receptors that are expressed on cells which are photoreceptors, in the eye. The recombinant virus comprises a fibre protein which is an adenovirus type 37, from an adenovirus type D serotype. The fibre is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fibre protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient portion of an adenovirus serotype D knob portion of the fiber for selective binding to photoreceptors in the eye of a mammal. The encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a nucleic acid comprising the therapeutic product which is chosen from trophic factor, anti-apoptotic factor, gene encoding a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-specific gene product. The delivery is effected for treatment of an ocular disease such as retinal degenerative disease e.g., retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or retinoblastoma inherited and acquired retinal and neovascular degenerative diseases. The viral nucleic acid comprises an adenovirus inverted terminal repeat (ITR) sequences, and an adenovirus packaging signal operatively linked to the sequence. The ITRs and packaging signal are derived from an adenovirus serotype B or C, or adenovirus type 2 or 5. The viral nucleic acid further comprises a photoreceptor-specific promoter. (II) includes photoreceptor promoters providing a means not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression. This sequence represents a TPL (tripartite leader sequence) from the adenovirus type 5 genome, used to enhance the expression of complementing adenoviral proteins.

Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;

Query Match 100.0%; Score 327; DB 10; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTTCTTT 60  
 DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTTCTTT 60  
 QY 61 CCAGTACTCTTGGATCGGAACCCGCTCGGCTCCGAAACGGTACTCCGCCACCGAGGACC 120  
 DB 61 CCAGTACTCTTGGATCGGAACCCGCTCGGCTCCGAAACGGTACTCCGCCACCGAGGACC 120  
 QY 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAAACGATCAC 180  
 DB 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGTCTAAACGATCAC 180  
 QY 181 AGTCGCAAGGTAGGCTGAGACCCGTCGGCGGGCGGAGCGGGTGGCGGTTGTTTC 240  
 DB 181 AGTCGCAAGGTAGGCTGAGACCCGTCGGCGGGCGGAGCGGGTGGCGGTTGTTTC 240  
 QY 241 TGGCGGAGGTCTGCTGATGATGTAATTAAAGTAGGCGGCTTTGAGACGGCGGATGTCG 300  
 DB 241 TGGCGGAGGTCTGCTGATGATGTAATTAAAGTAGGCGGCTTTGAGACGGCGGATGTCG 300  
 QY 301 AGGTGAGGTGTCGACGAGGCTTCAGATCT 327  
 DB 301 AGGTGAGGTGTCGACGAGGCTTCAGATCT 327

# RESULT 4

ADFA8736

ID ADF48736 standard; DNA; 327 BP.

XX ADF48736;

XX 12-FEB-2004 (first entry)

DE Adenovirus type 5 partial tripartite leader sequence.

XX cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;  
 KW HIV gene expression activation; adenovirus tripartite leader; TPL;  
 KW gutless adenoviral vector particle;  
 KW helper-independent fiberless recombinant adenovirus vector;  
 KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;  
 KW hereditary disorder; tumour; HIV infection; fibre;  
 KW fibre-gene-deleted adenoviruses; hygromycin resistance;  
 KW tripartite leader sequence; ds.

OS Human adenovirus type 5.

FN US2003157688-A1.

XX 21-AUG-2003.

PF 14-JAN-2000; 2000US-00482682.

PR 14-JAN-1999; 99US-0115920P.

PR 26-JUN-2000; 2000US-00423783.

XX (VSEG/) VON SEGGERN D J.

PA (NEME/) NEMEROW G R.

PA (HALL/) HALLENBECK P.

PA (STEV/) STEVENSON S.

PA (SKRI/) SKRIPCHENKO Y.

XX Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;

PI Skripchenko Y;

XX WPI; 2003-843463/78.

XX Novel isolated nucleic acid molecule useful for delivering heterologous gene to human or any animal, or for producing gutless adenoviral vector

PT particle.  
 XX Claim 14; SEQ ID NO 26; 157pp; English.  
 XX  
 CC The invention describes an isolated nucleic acid molecule (I) comprising  
 CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide  
 CC sequence comprising a first and second different TPL exons or first,  
 CC second and third same or different TPL exons, the TPL exons chosen from  
 CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon  
 CC 3. (I) is useful for delivering a heterologous gene to a human or any  
 CC animal, or for producing a gutless adenoviral vector particle. A  
 CC recombinant adenovirus particle (II) is useful for delivery of an  
 CC exogenous gene to a target cell which involves contacting the cell with  
 CC an amount of (II) sufficient to infect the cell. A helper-independent  
 CC fiberless recombinant adenovirus vector genome (III) is useful for  
 CC producing an adenovirus vector particle containing (III) which involves  
 CC providing a packaging cell line which complements replication and  
 CC packaging of the genome and (III) which is deficient in expressing  
 CC sufficient functional fiber protein to support assembly of fiber  
 CC containing particles and harvesting the particle produced by the cell  
 CC line. (III) is useful for pseudotyping recombinant viral vectors which  
 CC involves complementing a missing fiber gene of (III) or helper-dependent  
 CC fiberless recombinant adenovirus vector genome by expressing in packaging  
 CC cells a fiber gene from a different adenoviral serotype than the  
 CC recombinant adenovirus vector. (III) is also useful for specifically  
 CC targeting an adenovirus vector to a cell of choice. (I) is useful for  
 CC gene therapy. (II) is useful for treating diseases such as hereditary  
 CC disorder, and for reducing proliferation of tumour cells in a subject, or  
 CC to disrupt HIV infection. This sequence represents a partial adenovirus  
 CC tripartite leader sequence.  
 XX  
 SQ Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;

Query Match 100.0%; Score 327; DB 10; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTT 60  
 DB 1 AGATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTT 60  
 QY 61 CCACTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTCGCCACCGAGGGACC 120  
 DB 61 CCACTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTCGCCACCGAGGGACC 120  
 QY 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGGTCTAACCACTCAC 180  
 DB 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGGTCTAACCACTCAC 180  
 QY 181 AGTCGCAAGGTAGGCTGAGCACCGTGGCGGCGGCGGCGGCTGGCGGTGTTTC 240  
 DB 181 AGTCGCAAGGTAGGCTGAGCACCGTGGCGGCGGCGGCGGCTGGCGGTGTTTC 240  
 QY 241 TGGCGGAGTGTGCTGTGATGATGATTAATTAAGTAGCGGTCTTGAGACGGCGGATGGTCG 300  
 DB 241 TGGCGGAGTGTGCTGTGATGATGATTAATTAAGTAGCGGTCTTGAGACGGCGGATGGTCG 300  
 QY 301 AGGTGAGGTGTGGCGGCTTGAGATCT 327  
 DB 301 AGGTGAGGTGTGGCGGCTTGAGATCT 327

RESULT 5  
 AAV32372  
 ID AAV32372 standard; DNA; 7469 BP.

XX AAV32372;  
 AC  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 29-SEP-1998 (first entry)  
 XX  
 DE Complete sequence of the pCLF plasmid.

KW Circular; adenovirus type 5; pCDNA3/Fiber plasmid; structural protein;  
 KW complementation; fiber protein; gene therapy; HIV; tumour; early gene;  
 KW Huntington's disease; Tay-Sachs disease; sickle cell disease;  
 KW pCLF plasmid; AD2; adenovirus type 2; ds.  
 XX Synthetic.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT sig\_peptide 907..1233  
 FT /\*tag= a  
 FT /note= "AD2 tripartite leader sequence"  
 FT CDS 1237..2982  
 FT /\*tag= b  
 FT /product= "AD5 fiber protein"  
 XX  
 XX W09813499-A2.  
 XX  
 XX 02-APR-1998.  
 XX  
 XX 24-SEP-1997; 97WO-EP005251.  
 XX  
 XX 25-SEP-1996; 96US-00719806.  
 XX  
 XX (NOVS ) NOVARTIS AG.  
 XX (SCRI ) SCRIPPS RES INST.  
 XX  
 XX Nemerow GR, Von Seggern DJ;  
 XX WPI; 1998-230709/20.  
 XX  
 XX Adenoviral vectors - which lack DNA encoding for structural protein or  
 XX fibre protein used particularly for gene therapy.

Example 1; Page 85-94; 170pp; English.

The present sequence is that of a pCLF plasmid used in the method of the invention. The pCLF plasmid was derived from the pCDNA3/Fiber plasmid (AAV32371) containing an additional adenovirus type 2 (AD2) tripartite leader sequence to enhance expression. The pCLF plasmid also contains an adenovirus type 5 (AD5) fiber gene controlled by a CMV promoter and a neo resistant gene. The invention provides adenoviral vectors having deletions of all or part of various gene sequences encoding adenoviral structural proteins and/or early region proteins. Deletions in these proteins would allow a reduced risk of wild-type virus contamination and would also allow packaging of foreign DNA in such vectors for a variety of diagnostic and therapeutic applications. The adenoviral vectors having deletions in the structural and/or early gene regions are produced by cellular complementation of these adenoviral genes. Therefore, the pCLF plasmid was used as a complementation plasmid which was introduced into a host cell line where parts of the fiber gene region would be stably inserted into the host cell chromosomes. The resulting fiber gene deficient plasmid can be used as a gene delivery vector. The vectors can be used for diagnosis or gene therapy, e.g. for treating conditions characterised by hyper-proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g. HIV infection). They can also be used for in vitro production of biologically active proteins. (Updated on 25-MAR-2003 to correct PI field.)

Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 2; Length 7469;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-79;  
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GATCTGAATTCGAGTCTGCTGTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTTC 61  
 DB 908 GATCTGAATTCGAGTCTGCTGTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTTC 967  
 QY 62 CAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTTCGCCACCGAGGACCT 121  
 DB 968 CAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTTCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTCGAGAAAGCGGTCTAACAGTCACA 181  
 Db 1028 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTCGAGAAAGCGGTCTAACAGTCACA 1087  
 QY 182 GTCGCAAGGTAGCTGAGCACCGTGGCGCGCGGCGGAGCGGTGGCGGTCTTCT 241  
 Db 1088 GTCGCAAGGTAGCTGAGCACCGTGGCGCGCGGCGGAGCGGTGGCGGTCTTCT 1147  
 QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 301  
 Db 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 1207  
 QY 302 GGTGAGGTGTCAGGCTTGAGATC 326  
 Db 1208 GGTGAGGTGTCAGGCTTGAGATC 1232

## RESULT 6

AAAS9043  
 ID AAAS9043 standard; DNA; 7469 BP.

AC AAAS9043;

DT 15-SEP-2003 (revised)

DT 07-NOV-2000 (first entry)

DE Nucleotide sequence of a partial tripartite leader sequence.

KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;  
 KW ss.

OS Human adenovirus type 5.

XX WO200042208-A1.

XX 20-JUL-2000.

PF 14-JAN-2000; 2000WO-EP000265.

PR 14-JAN-1999; 99US-0115920P.

XX (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA (SCRI ) SCRIPPS RES INST.

PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;

PI Skripchenko Y;

XX WPI; 2000-476068/41.

XX New nucleic acid comprising an adenovirus tripartite leader nucleotide  
 PT for producing high-capacity and targeted vectors for adenovirus-based  
 PT gene therapy.

PS Claim 10; Page 154-156; 212pp; English.

XX The specification describes a nucleic acid molecule comprising an  
 CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence  
 CC comprising two different TPL exons or three same or different TPL exons.  
 CC The nucleic acid is used to produce an adenovirus vector particle,  
 CC deliver an exogenous gene to a target cell, pseudotype recombinant viral  
 CC vectors, target an adenovirus vector to a cell, produce a modified  
 CC adenovirus, deliver a heterologous gene to an animal and produce a  
 CC gutless adenoviral vector particle. The present sequence represents a  
 CC partial TPL sequence, which is used to construct nucleic acid molecules  
 CC of the invention. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 3; Length 7469;

Best Local Similarity 100.0%; Pred. No. 3.7e-79;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGGCTCGCGGTGAGCAAACTCTTCGGGTCTTTTC 61  
 Db 908 GATCTGAATTCGAGCTCGCTGTGGGCTCGCGGTGAGCAAACTCTTCGGGTCTTTTC 967  
 QY 62 CAGTACTCTTCGATCGGAAACCGCTCGGCTCCGACCGTACTCCGCCACCGAGGACCT 121  
 Db 968 CAGTACTCTTCGATCGGAAACCGCTCGGCTCCGACCGTACTCCGCCACCGAGGACCT 1027  
 QY 122 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTTCGAGAAAGCGGTCTAACAGTCACA 181  
 Db 1028 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTTCGAGAAAGCGGTCTAACAGTCACA 1087  
 QY 182 GTCGCAAGGTAGCTGAGCACCGTGGCGCGCGGCGGAGCGGTGGCGGTCTTCT 241  
 Db 1088 GTCGCAAGGTAGCTGAGCACCGTGGCGCGCGGCGGAGCGGTGGCGGTCTTCT 1147  
 QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 301  
 Db 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 1207  
 QY 302 GGTGAGGTGTCAGGCTTGAGATC 326  
 Db 1208 GGTGAGGTGTCAGGCTTGAGATC 1232

## RESULT 7

ABA94252

XX ID ABA94252 standard; DNA; 7469 BP.

AC ABA94252;

DT 13-MAR-2002 (first entry)

DE Nucleotide sequence of expression plasmid pCLF.

XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
 KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;  
 KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
 KW gene therapy; ss.

OS Synthetic.

XX WO200183729-A2.

XX 08-NOV-2001.

XX 30-APR-2001; 2001WO-EP004863.

XX 01-MAY-2000; 2000US-00562934.

XX (NOVS ) NOVARTIS AG.

PA (SCRI ) SCRIPPS RES INST.

PA (NEME/) NEMEROW G R.

PA (VSEG/) VON SEGGERN D J.

PA (FRIE/) FRIEDLANDER M.

PI Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2002-082846/11.

XX Polynucleotide for making vectors, useful for treating ocular diseases,  
 PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat  
 PT sequences, packaging signal and photoreceptor-specific promoter.

XX Example 1; Page 108-110; 149pp; English.

XX The invention provides an isolated polynucleotide comprising adenovirus  
 CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal  
 CC operatively linked to ITRS and a photoreceptor-specific promoter. A  
 CC recombinant AV vector (AVV) comprising the polynucleotide is useful for  
 CC targeted delivery of a gene product to the eye (especially to the  
 CC vitreous cavity), for treating an ocular disease, e.g., retinal  
 CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic



Db 968 CAGTACTCTTGGATCGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 1027  
QY 122 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGCGGTCTTAACAGTCACA 181  
Db 1028 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGCGGTCTTAACAGTCACA 1087  
QY 182 GTCCGAAGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTGGCGGTGGCGGT 241  
Db 1088 GTCCGAAGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTGGCGGTGGCGGT 1147  
QY 242 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGTCGA 301  
Db 1148 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGTCGA 1207  
QY 302 GGTGAGGTGGCAGGCTTGAGATC 326  
Db 1208 GGTGAGGTGGCAGGCTTGAGATC 1232  
RESULT 9  
ID ADF48794  
XX ADF48794 standard; DNA; 7469 BP.  
AC ADF48794;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Fibre expressing/tripartite leader sequence plasmid pCLF.  
XX  
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;  
KW HIV gene expression activation; adenovirus tripartite leader; TPL;  
KW gutless adenoviral vector particle;  
KW helper-independent fiberless recombinant adenovirus vector;  
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;  
KW hereditary disorder; tumour; HIV infection; fibre;  
KW fibre-gene-deleted adenoviruses; hygromycin resistance;  
KW tripartite leader sequence; ds; pCLF; pCDN3/fibre.  
XX  
OS Human adenovirus type 5.  
XX  
FN US2003157688-A1.  
XX  
PD 21-AUG-2003.  
XX  
PF 14-JAN-2000; 2000US-00482682.  
XX  
PR 14-JAN-1999; 99US-0115920P.  
PR 26-JUN-2000; 2000US-00423783.  
XX  
PA (VSEG/) VON SEGGERN D J.  
PA (NEME/) NEMEROW G R.  
PA (HALL/) HALLENBECK P.  
PA (STEV/) STEVENSON S.  
PA (SKRI/) SKRIPCHENKO Y.  
XX  
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;  
PI Skripchenko Y;  
XX  
DR WPI; 2003-843463/78.  
XX  
PT Novel isolated nucleic acid molecule useful for delivering heterologous  
PT gene to human or any animal, or for producing gutless adenoviral vector  
PT particle.  
XX  
PS Claim 10; SEQ ID NO 8; 157pp; English.  
XX  
CC The invention describes an isolated nucleic acid molecule (I) comprising  
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide  
CC sequence comprising a first and second different TPL exons or first,  
CC second and third same or different TPL exons, the TPL exons chosen from  
CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon  
CC 3. (I) is useful for delivering a heterologous gene to a human or any  
CC animal, or for producing a gutless adenoviral vector particle. A.

CC recombinant adenovirus particle (II) is useful for delivery of an  
CC exogenous gene to a target cell which involves contacting the cell with  
CC an amount of (II) sufficient to infect the cell. A helper-independent  
CC fiberless recombinant adenovirus vector genome (III) is useful for  
CC producing an adenovirus vector particle containing (III) which involves  
CC providing a packaging cell line which complements replication and  
CC packaging of the genome and (III) which is deficient in expressing  
CC sufficient functional fiber protein to support assembly of fiber  
CC containing particles and harvesting the particle produced by the cell  
CC line. (III) is useful for pseudotyping recombinant viral vectors which  
CC involves complementing a missing fiber gene of (III) or helper-dependent  
CC fiberless recombinant adenovirus vector genome by expressing in packaging  
CC cells a fiber gene from a different adenoviral serotype than the  
CC recombinant adenovirus vector. (III) is also useful for specifically  
CC targeting an adenovirus vector to a cell of choice. (I) is useful for  
CC gene therapy. (II) is useful for treating diseases such as hereditary  
CC disorder, and for reducing proliferation of tumour cells in a subject, or  
CC to disrupt HIV infection. This sequence represents an adenovirus  
CC tripartite leader sequence added to plasmid pCDN3/fibre to create plasmid  
CC pCLF, an adenovirus fibre expressing plasmid for complementation of E4-  
CC gene-deleted adenoviruses.  
XX  
SQ Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;  
Query Match 99.4%; Score 325; DB 10; Length 7469;  
Best Local Similarity 100.0%; Pred. No. 3.7e-79;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GATCTGAATTCGAGTCGCTCTTGGGCTCGCGGTGAGGACAACTCTTCGGGCTTTTC 61  
Db 908 GATCTGAATTCGAGTCGCTCTTGGGCTCGCGGTGAGGACAACTCTTCGGGCTTTTC 967  
QY 62 CAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121  
Db 968 CAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 1027  
QY 122 GAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGGCGTCTAACCACTACA 181  
Db 1028 GAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGGCGTCTAACCACTACA 1087  
QY 182 GTCCGAAGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTGGCGGTGGCGGT 241  
Db 1088 GTCCGAAGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTGGCGGTGGCGGT 1147  
QY 242 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGTCGA 301  
Db 1148 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGTCGA 1207  
QY 302 GGTGAGGTGGCAGGCTTGAGATC 326  
Db 1208 GGTGAGGTGGCAGGCTTGAGATC 1232  
RESULT 10  
AAV32375  
ID AAV32375 standard; DNA; 10610 BP.  
XX  
AC AAV32375;  
XX  
DT 25-MAR-2003 (revised)  
DT 29-SEP-1998 (first entry)  
XX  
DE Complete sequence of the pE4/Fiber plasmid.  
XX  
KW Circular; adenovirus type 5; pE4/Fiber plasmid; structural protein;  
KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;  
KW early gene; Huntington's disease; Tay-Sachs disease; sickle cell disease;  
KW E4 regulatory gene; ds.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature complement(21..3149)

```
FT      /*tag= a
FT      /note= "AD5 E4 regulatory gene"
FT      4051..4366
FT      /*tag= b
FT      /note= "AD5 leader sequence"
FT      4372..6124
FT      /*tag= c
FT      /note= "AD5 fiber gene"
XX
PN      WO9813499-A2.
XX
XX      02-APR-1998.
XX
XX      24-SEP-1997; 97WO-EP005251.
XX
XX      25-SEP-1996; 96US-00719806.
XX
XX      (NOVS ) NOVARTIS AG.
XX      (SCRI ) SCHIPPS RES INST.
XX
PI      Nemerow GR, Von Seggern DJ;
XX      WPI; 1998-230709/20.
XX
XX      Adenoviral vectors - which lack DNA encoding for structural protein or
XX      fibre protein used particularly for gene therapy.
XX
XX      Example 1; Page 131-145; 170pp; English.
XX
XX      The present sequence is that of a pE4/Fiber plasmid used in the method of
XX      the invention. The plasmid contains an adenovirus type 5 (AD5) fiber gene
XX      controlled by a CMV promoter, an AD5 E4 gene and an adenovirus type 2
XX      (AD2) tripartite leader sequence upstream of the fiber gene. The
XX      invention provides adenoviral vectors having deletions of all or part of
XX      various gene sequences encoding adenoviral structural proteins and/or
XX      early region proteins. Deletions in these proteins would allow a reduced
XX      risk of wild-type virus contamination and would also allow packaging of
XX      foreign DNA in such vectors for a variety of diagnostic and therapeutic
XX      applications. The adenoviral vectors having deletions in the structural
XX      and/or early gene regions are produced by cellular complementation of
XX      these adenoviral genes. Therefore, the pE4/Fiber plasmid was used as a
XX      complementation plasmid which was introduced into a host cell line where
XX      parts of the fiber and E4 gene region would be stably inserted into the
XX      host cell chromosomes. The resulting E4/fiber gene deficient plasmid can
XX      be used as a gene delivery vector. The vectors can be used for diagnosis
XX      or gene therapy, e.g. for treating conditions characterised by hyper-
XX      proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's
XX      disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g.
XX      HIV infection). They can also be used for in vitro production of
XX      biologically active proteins. (Updated on 25-MAR-2003 to correct PI
XX      field.)
XX
XX      Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;
XX
XX      Query Match      99.4%; Score 325; DB 2; Length 10610;
XX      Best Local Similarity 100.0%; Pred. No. 4e-79;
XX      Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      2 GATCTGAATTCGAGTCGCTGTGGGCTCGCGGTTGAGGACAAACTCTTCGGCGTCTTTC 61
XX      4049 GATCTGAATTCGAGTCGCTGTGGGCTCGCGGTTGAGGACAAACTCTTCGGCGTCTTTC 4108
XX
XX      62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
XX      4109 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168
XX
XX      122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGCGGTCTAACCACTCACA 181
XX      4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGCGGTCTAACCACTCACA 4228
XX
XX      182 GTCCGAAGTAGGCTAGCAGCACCGTCGGCGGGCGGACGGGTGGCGGTTGTTTCT 241
XX      4229 GTCCGAAGTAGGCTAGCAGCACCGTCGGCGGGCGGACGGGTGGCGGTTGTTTCT 4288
```

```
QY      242 GCGGAGGTGCTGCTGATGATGTAATTAAAGTAGCGGCTTTGAGACGGCGGATGCTCGA 301
Db      4289 GCGGAGGTGCTGCTGATGATGTAATTAAAGTAGCGGCTTTGAGACGGCGGATGCTCGA 4348
QY      302 GGTGAGGTGTGGCAGGCTTGAGATC 326
Db      4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 11
AAAS9051
ID      AAAS9051 standard; DNA; 10610 BP.
XX
XX      AC      AAAS9051;
XX
XX      DT      07-NOV-2000 (first entry)
XX
XX      DE      Nucleotide sequence of the E4/fiber-expressing plasmid pE4/Fiber.
XX
XX      KW      Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
XX      E4 gene; fiber gene; ss.
XX
XX      OS      Synthetic.
XX      QS      Human adenovirus type 5.
XX      XX      WO200042208-A1.
XX      PD      20-JUL-2000.
XX
XX      PF      14-JAN-2000; 2000WO-EP000265.
XX
XX      PR      14-JAN-1999; 99US-0115920P.
XX
XX      PA      (NOVS ) NOVARTIS AG.
XX      (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX      (SCRI ) SCHIPPS RES INST.
XX
XX      PI      Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
XX      Skripchenko Y;
XX      WPI; 2000-476068/41.
XX
XX      PT      New nucleic acid comprising an adenovirus tripartite leader nucleotide
XX      for producing high-capacity and targeted vectors for adenovirus-based
XX      gene therapy.
XX
XX      PS      Example 1; Page 164-167; 212pp; English.
XX
XX      CC      The specification describes a nucleic acid molecule comprising an
XX      adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
XX      comprising two different TPL exons or three same or different TPL exons.
XX      The nucleic acid is used to produce an adenovirus vector particle,
XX      deliver an exogenous gene to a target cell, pseudotype recombinant viral
XX      vectors, target an adenovirus vector to a cell, produce a modified
XX      adenovirus, deliver a heterologous gene to an animal and produce a
XX      gutless adenoviral vector particle. The present sequence represents
XX      pE4/Fiber, a complementing plasmid containing E4 and fiber Adenoviral
XX      genes
XX
XX      SQ      Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;
XX
XX      Query Match      99.4%; Score 325; DB 3; Length 10610;
XX      Best Local Similarity 100.0%; Pred. No. 4e-79;
XX      Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      2 GATCTGAATTCGAGTCGCTGTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 61
XX      4049 GATCTGAATTCGAGTCGCTGTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 4108
XX
XX      62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
XX      4109 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168
```

QY 122 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAAGAGGGCTTAACCACTCACA 181  
DB 4169 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAAGAGGGCTTAACCACTCACA 4228  
QY 182 GTCCGAAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGCTGGGGTTGTTTCT 241  
DB 4229 GTCCGAAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGCTGGGGTTGTTTCT 4288  
QY 242 GCGCGAGGTGCTGCTGATGATGTAATAAGTAGGCGGTCTTGAGACGCGGATGCTCGA 301  
DB 4289 GCGCGAGGTGCTGCTGATGATGTAATAAGTAGGCGGTCTTGAGACGCGGATGCTCGA 4348  
QY 302 GGTGAGGTGTGCGAGGCTTGAGATC 326  
DB 4349 GGTGAGGTGTGCGAGGCTTGAGATC 4373

RESULT 12  
ABA94260  
ID ABA94260 standard; DNA; 10610 BP.  
AC ABA94260;  
XX  
XX  
XX  
DT 13-MAR-2002 (first entry)  
XX  
DE Nucleotide sequence of expression plasmid pB4/Fiber.  
XX  
KW Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;  
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; ss.  
XX  
OS Synthetic.  
XX  
XX WO200183729-A2.  
PN  
PD 08-NOV-2001.  
XX  
XX 30-APR-2001; 2001WO-EP004863.  
XX  
XX 01-MAY-2000; 2000US-00562934.  
XX  
XX (NOVS ) NOVARTIS AG.  
XX (SCRI ) SCRIPPS RES INST.  
XX (NEME/) NEMEROW G R.  
XX (VSEG/) VON SEGGERN D J.  
XX (FRIE/) FRIEDLANDER M.  
XX  
XX Nemerow GR, Von Seggern DJ, Friedlander M;  
XX WPI; 2002-082846/11.  
XX  
XX Polynucleotide for making vectors, useful for treating ocular diseases,  
XX e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat  
XX sequences, packaging signal and photoreceptor-specific promoter.  
XX  
XX Example 1; Page 118-121; 149pp; English.  
XX  
XX The invention provides an isolated polynucleotide comprising adenovirus  
XX (AV) inverter terminal repeat sequences (ITRS), AV packaging signal  
XX operatively linked to ITRS and a photoreceptor-specific promoter. A  
XX recombinant AV vector (AVV) comprising the polynucleotide is useful for  
XX targeted delivery of a gene product to the eye (especially to the  
XX vitreous cavity), for treating an ocular disease, e.g., retinal  
XX degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic  
XX retinopathies, retinal vascularizations, and retinoblastoma, of a mammal  
XX preferably human. The AAV comprises a fiber protein that specifically or  
XX selectively binds to receptors that are expressed on cells (preferably  
XX photoreceptors in the eye). Preferably, the recombinant virus comprise a  
XX fiber protein from an adenovirus type D subgroup or is a chimeric protein  
XX containing a portion of the N-terminus of an adenovirus type 2 or type 5  
XX penton, and the therapeutic product is a trophic factor, an anti-

CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type  
CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that  
CC regulates expression of a photoreceptor specific gene product. The viral  
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV  
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful  
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber  
CC protein or its portion, and selectively transduces photoreceptors and  
CC delivers a gene product encoded by AAV. The present sequence represents  
CC an expression plasmid pB4/Fiber containing the adenovirus E4 and Fiber  
CC genes  
XX  
XX Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;  
SQ  
Query Match 99.4%; Score 325; DB 6; Length 10610;  
Best Local Similarity 100.0%; Pred. No. 4e-79;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTGAGGACAAACTCTTCGGGCTTTTC 61  
DB 4049 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTGAGGACAAACTCTTCGGGCTTTTC 4108  
QY 62 CAGTACTCTTGGATCGGAACCCCGTCGGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 121  
DB 4109 CAGTACTCTTGGATCGGAACCCCGTCGGCCCTCCGAACGGTACTCCGCCACCGAGGACCT 4168  
QY 122 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAAGAGGGCTTAACCACTCACA 181  
DB 4169 GAGCGAGTCCGATCGACCGGATCGAAACCTCTCGAAGAGGGCTTAACCACTCACA 4228  
QY 182 GTCCGAAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGCTGGGGTTGTTTCT 241  
DB 4229 GTCCGAAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGCTGGGGTTGTTTCT 4288  
QY 242 GCGCGAGGTGCTGCTGATGATGTAATAAGTAGGCGGTCTTGAGACGCGGATGCTCGA 301  
DB 4289 GCGCGAGGTGCTGCTGATGATGTAATAAGTAGGCGGTCTTGAGACGCGGATGCTCGA 4348  
QY 302 GGTGAGGTGTGCGAGGCTTGAGATC 326  
DB 4349 GGTGAGGTGTGCGAGGCTTGAGATC 4373  
RESULT 13  
ADB75157  
ID ADB75157 standard; DNA; 10610 BP.  
XX  
XX ADB75157;  
AC  
XX  
XX  
DT 04-DEC-2003 (first entry)  
XX  
XX Plasmid pB4/Fibre DNA sequence.  
XX  
XX ophthalmological; antiinflammatory; antidiabetic; gene therapy;  
KW adenovirus inverted terminal repeat sequence;  
KW adenovirus packaging signal; photoreceptor-specific promoter;  
KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;  
KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;  
KW rhodopsin; wild-type Stargardt disease gene; STDG1; anti-cancer agent;  
KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;  
KW gyrate atrophy; macular dystrophy; retinoblastoma;  
KW photoreceptor-restricted transgene expression;  
KW recombinant adenovirus vector; adenovirus type 5; pB4/Fibre; plasmid;  
KW cyclic; circular; dB; B4; fibre.  
XX  
XX Synthetic.  
OS  
XX Human adenovirus type 5.  
XX  
XX US2002193327-A1.  
FN  
XX  
XX 19-DEC-2002.  
PD  
XX  
XX 01-MAY-2001; 2001US-00847101.  
PF

XX 01-MAY-2000; 2000US-00562934.  
PR (SCRI ) SCHRIPPS RES INST.  
PA Nemerow GR, Von Seggern DJ, Friedlander M;  
PI WPI; 2003-657234/62.  
XX  
DR  
XX  
XX Novel nucleic acids comprising adenovirus inverted terminal repeat  
PT sequences, adenovirus packaging signals operatively linked to the  
PT sequences and photoreceptor-specific promoters, useful for treating  
PT retinitis pigmentosa.  
XX  
PS Example 1; Page 57-61; 106pp; English.  
XX  
CC The invention describes an isolated nucleic acid (I) comprising  
CC adenovirus inverted terminal repeat sequence, an adenovirus packaging  
CC signal operatively linked to the sequence, and a photoreceptor-specific  
CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful  
CC for targeted delivery of a gene product to the eye of a mammal which  
CC involves administering (II) that comprises heterologous DNA encoding the  
CC gene product or resulting in expression of the gene product, where the  
CC recombinant virus comprises a fibre protein that specifically or  
CC selectively binds to receptors that are expressed on cells which are  
CC photoreceptors, in the eye. The recombinant virus comprises a fibre  
CC protein which is an adenovirus type 37, from an adenovirus type D  
CC serotype. The fibre is a chimeric protein containing a sufficient portion  
CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for  
CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient  
CC portion of an adenovirus serotype D knob portion of the fiber for  
CC selective binding to photoreceptors in the eye of a mammal. The  
CC encapsulated nucleic acid comprises a photoreceptor-specific promoter  
CC operatively linked to a nucleic acid comprising the therapeutic product  
CC which is chosen from tropic factor, anti-apoptotic factor, gene encoding  
CC a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-  
CC cancer agent and a protein that regulates expression of a photoreceptor-  
CC specific gene product. The delivery is effected for treatment of an  
CC ocular disease such as retinal degenerative disease e.g., retinitis  
CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal  
CC vascularisation, choroidaemia, gyrate atrophy or macular dystrophy or  
CC retinoblastoma inherited and acquired retinal and neovascular  
CC degenerative diseases. The viral nucleic acid comprises an adenovirus  
CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging  
CC signal operatively linked to the sequence. The ITRs and packaging signal  
CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or  
CC 5. The viral nucleic acid further comprises a photoreceptor-specific  
CC promoter. (II) includes photoreceptor promoters providing a means not  
CC only for specific targeting of expression in these cells, but also for  
CC photoreceptor-restricted transgene expression. This sequence represents a  
CC plasmid expressing adenovirus type 5 fibre gene and E4 gene that can be  
CC used to complement one or more delivery plasmids expressing E4 and fibre.  
SQ Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;  
XX

Query Match 99.4%; Score 325; DB 10; Length 10610;  
Best Local Similarity 100.0%; Pred. No. 4e-79;  
Matches: 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 2 GATCTGAATTCGAGTTCGGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGTCCTTC 61  
DB 4049 GATCTGAATTCGAGTTCGGCTGTGGCTCGCGGTTGAGACAACTCTTCGGGTCCTTC 4108  
QY 62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGACCGGTACTCCGCCACCGAGGACCT 121  
DB 4109 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGACCGGTACTCCGCCACCGAGGACCT 4168  
QY 122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTTCGAGAAAGCGCTTCAACCACTACA 181  
DB 4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTTCGAGAAAGCGCTTCAACCACTACA 4228  
QY 182 GTCCCAAGTAGCTGAGCACCCTGGCGGGCGGACGCGGTGGCGGTTCGGGTCTTTCT 241  
XX

DB 4229 GTCCCAAGTAGCTGAGCACCCTGGCGGGCGGACGGGTGGCGGTTCCTTC 4288  
QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTACGGCGGTCTTGAGACGCGGATGGTCGA 301  
DB 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTACGGCGGTCTTGAGACGCGGATGGTCGA 4348  
QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326  
DB 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373  
RESULT 14  
ADF48802  
ID ADF48802 standard; DNA; 10610 BP.  
XX  
AC ADF48802;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE E4/fibre expressing plasmid pE1/fibre.  
XX  
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;  
KW HIV gene expression activation; adenovirus tripartite leader; TPL;  
KW gutless adenoviral vector particle;  
KW helper-independent fiberless recombinant adenovirus vector;  
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;  
KW hereditary disorder; tumour; HIV infection; E4 transcription unit; fibre;  
KW hygromycin resistance; ds; circular; cyclic.  
OS Synthetic.  
OS Human adenovirus type 5.  
FN US2003157688-A1.  
XX  
PD 21-AUG-2003.  
XX  
PF 14-JAN-2000; 2000US-00482682.  
XX  
PR 14-JAN-1999; 99US-0115920P.  
PR 26-JUN-2000; 2000US-00423783.  
XX  
PA (VSEG//) VON SEGGERN D J.  
PA (NEME//) NEMEROW G R.  
PA (HALL//) HALLENBECK P.  
PA (STEV//) STEVENSON S.  
PA (SKRI//) SKRIPCHENKO Y.  
XX  
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;  
PI Skripchenko Y;  
XX  
DR WPI; 2003-843463/78.  
XX  
PT Novel isolated nucleic acid molecule useful for delivering heterologous  
PT gene to human or any animal, or for producing gutless adenoviral vector  
PT particle.  
XX  
PS Example 1; SEQ ID NO 16; 157pp; English.  
XX  
CC The invention describes an isolated nucleic acid molecule (I) comprising  
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide  
CC sequence comprising a first and second different TPL exons or first  
CC second and third same or different TPL exons, the TPL exons chosen from  
CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon  
CC 3. (I) is useful for delivering a heterologous gene to a human or any  
CC animal, or for producing a gutless adenoviral vector particle. A  
CC recombinant adenovirus particle (II) is useful for delivery of an  
CC exogenous gene to a target cell which involves contacting the cell with  
CC an amount of (II) sufficient to infect the cell. A helper-independent  
CC fiberless recombinant adenovirus vector genome (III) is useful for  
CC producing an adenovirus vector particle containing (III) which involves  
CC providing a packaging cell line which complements replication and  
CC packaging of the genome and (III) which is deficient in expressing  
CC sufficient functional fiber protein to support assembly of fiber



CC containing particles and harvesting the particle produced by the cell  
 CC line. (III) is useful for pseudotyping recombinant viral vectors which  
 CC involves complementing a missing fiber gene of (III) or helper-dependent  
 CC fiberless recombinant adenovirus vector genome by expressing in packaging  
 CC cells a fiber gene from a different adenoviral serotype than the  
 CC recombinant adenovirus vector. (III) is also useful for specifically  
 CC targeting an adenovirus vector to a cell of choice. (I) is useful for  
 CC gene therapy. (II) is useful for treating diseases such as hereditary  
 CC disorder, and for reducing proliferation of tumour cells in a subject, or  
 CC to disrupt HIV infection. This sequence represents the complementing  
 CC plasmid pE4/fibre that expresses the adenoviral E4 and fibre genes.  
 XX  
 SQ Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 10; Length 10610;  
 Best Local Similarity 100.0%; Pred. No. 4e-79;  
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGTTGAGACAAACTCTTCGCGTCTTTC 61  
 Db 4049 GATCTGAATTCGAGTCGCTGTTGGCTCGCGTTGAGACAAACTCTTCGCGTCTTTC 4108  
 QY 62 CAGTACTCTTGGATCGGAACCGCTCGGCTCGAACGGTACTCCGCCACCGAGGACCT 121  
 Db 4109 CAGTACTCTTGGATCGGAACCGCTCGGCTCGAACGGTACTCCGCCACCGAGGACCT 4168  
 QY 122 GAGCAGTCCGATCGACCGATCGAAACCTCTCGAGAAAGGCTCTAACCACTACA 181  
 Db 4169 GAGCAGTCCGATCGACCGATCGAAACCTCTCGAGAAAGGCTCTAACCACTACA 4228  
 QY 182 GTCCAAAGTAGGCTGAGCACCGTGGCGGCGGACGCGGTGGCGGTGGCGGTCTTTC 241  
 Db 4229 GTCCAAAGTAGGCTGAGCACCGTGGCGGCGGACGCGGTGGCGGTGGCGGTCTTTC 4288  
 QY 242 GCGCAGGTGCTGCTGATGATCTAATAAGTAGGCGGTCTTGACACCGCGATGCTGA 301  
 Db 4289 GCGCAGGTGCTGCTGATGATCTAATAAGTAGGCGGTCTTGACACCGCGATGCTGA 4348  
 QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326  
 Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 15  
 AAV32374/c  
 ID AAV22374 standard; DNA; 14455 BP.

XX AAV32374;  
 AC

XX 25-MAR-2003 (revised)  
 DT

XX 29-SEP-1998 (first entry)  
 DT

XX Complete sequence of the pE1/Fiber plasmid.  
 DE

XX Circular; adenovirus type 5; pE1/Fiber plasmid; structural protein;  
 KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;  
 KW early gene; Huntington's disease; Tay-Sachs disease; sickle cell disease;  
 KW E1 regulatory protein; ds.  
 XX

OS Synthetic.  
 XX

XX Key Location/Qualifiers  
 FH 1460. .4998  
 FT misc\_feature  
 FT /tag= a

FT /note= "AD5 E1 regulatory gene"  
 FT complement (10922. .14223)  
 FT /tag= b

FT /note= "AD5 fiber gene consisting of a CMV promoter at 5'  
 FT end of this gene"  
 FT  
 XX

PN WO9813499-A2.  
 XX

XX 02-APR-1998.  
 PD

XX 24-SEP-1997; 97WO-EP005251.  
 PF  
 XX 25-SEP-1996; 96US-00719806.  
 PR  
 XX (NOVS ) NOVARTIS AG.  
 PA (SCKI ) SCRIPPS RES INST.  
 XX  
 PI Nemerow GR, Von Seggern DJ;  
 XX  
 XX WPI; 1998-230709/20.  
 XX  
 PT Adenoviral vectors - which lack DNA encoding for structural protein or  
 FT fibre protein used particularly for gene therapy.  
 XX  
 PS Example 1; Page 112-131; 170pp; English.  
 XX  
 CC The present sequence is that of a pE1/Fiber plasmid used in the method of  
 CC the invention. The plasmid contains an adenovirus type 5 (AD5) fiber gene  
 CC controlled by a CMV promoter, an AD5 E1 gene and a pMAM backbone. The  
 CC invention provides adenoviral vectors having deletions of all or part of  
 CC various gene sequences encoding adenoviral structural proteins and/or  
 CC early region proteins. Deletions in these proteins would allow a reduced  
 CC risk of wild-type virus contamination and would also allow packaging of  
 CC foreign DNA in such vectors for a variety of diagnostic and therapeutic  
 CC applications. The adenoviral vectors having deletions in the structural  
 CC and/or early gene regions are produced by cellular complementation of  
 CC these adenoviral genes. Therefore, the pE1/Fiber plasmid was used as a  
 CC complementation plasmid which was introduced into a host cell line where  
 CC parts of the fiber and E1 gene region would be stably inserted into the  
 CC host cell chromosomes. The resulting E1/fiber gene deficient plasmid can  
 CC be used as a gene delivery vector. The vectors can be used for diagnosis  
 CC or gene therapy, e.g. for treating conditions characterised by hyper-  
 CC proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's  
 CC disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g.  
 CC HIV infection). They can also be used for in vitro production of  
 CC biologically active proteins. (Updated on 25-MAR-2003 to correct PI  
 CC field.)  
 XX  
 SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 2; Length 14455;

Best Local Similarity 100.0%; Pred. No. 4.4e-79;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGTTGAGACAAACTCTTCGCGTCTTTC 61

Db 13315 GATCTGAATTCGAGTCGCTGTTGGCTCGCGTTGAGACAAACTCTTCGCGTCTTTC 13256

QY 62 CAGTACTCTTGGATCGGAACCGCTCGGCTCTCGAACGGTACTCCGCCACCGAGGACCT 121

Db 13255 CAGTACTCTTGGATCGGAACCGCTCGGCTCTCGAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCTCTAACCACTACA 181

Db 13195 GAGCAGTCCGATCGACCGGATCGAAACCTCTCGAGAAAGGCTCTAACCACTACA 13136

QY 182 GTCCAAAGTAGGCTGAGCACCGTGGCGGCGGACGCGGTGGCGGTGGCGGTCTTTC 241

Db 13135 GTCCAAAGTAGGCTGAGCACCGTGGCGGCGGACGCGGTGGCGGTGGCGGTCTTTC 13076

QY 242 GCGCAGGTGCTGCTGATGATCTAATAAGTAGGCGGTCTTGAGACCGCGATGCTGA 301

Db 13075 GCGCAGGTGCTGCTGATGATCTAATAAGTAGGCGGTCTTGAGACCGCGATGCTGA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326

Db 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

Search completed: July 14, 2005, 07:01:34

Job time : 465.082 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 3113.52 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-26

Perfect score: 327

Sequence: 1 agatctgaattcgatcgctgc.....gtgtgcaggcttgagatct 327

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	43.2	13.2	1086	9	CL467818 SAIL 1276
C 2	42.6	13.0	1262	9	CL496190 SAIL_620
C 3	41.6	12.7	1164	8	CC258768 CH261-164
C 4	41.6	12.7	1164	8	CC258769 CH261-164
C 5	40.4	12.4	710	4	BF972939 602241236
C 6	40.2	12.3	364	1	AJ602111 AJ602111
C 7	40.2	12.3	632	6	CA731965 wlpic.pk0
C 8	40	12.2	598	6	CA378248 657011 NC
C 9	40	12.2	1177	8	CC286173 CH261-29F
C 10	39.6	12.1	936	9	CL513605 SAIL 877
C 11	39.2	12.0	511	1	AJ437840 AJ437840
C 12	39.2	12.0	629	6	CA728476 wdiic.pk0
C 13	39.2	12.0	719	8	BZ715587 QGEAT34TC
C 14	39.2	12.0	847	5	EX900296 BX900296
C 15	39	11.9	378	6	CB685324 OSUNEF15K
C 16	39	11.9	608	6	CB648908 OSUNEB121
C 17	39	11.9	652	6	CB665091 OSUNEB11D
C 18	39	11.9	703	6	CB662732 OSUNED07A
C 19	39	11.9	726	9	AG101098 Pan trogl
C 20	39	11.9	831	7	CF711474 CCAD883TR
C 21	39	11.9	885	9	CC697699 OGNAJ58TH
C 22	39	11.9	1224	9	CL974456 OsIFCC025
C 23	38.6	11.8	384	1	AJ602280 AJ602280
C 24	38.6	11.8	621	6	CA733511 wlpic.pk0

C 25	38.6	11.8	622	4	BU280072
C 26	38.6	11.8	649	6	CD875521
C 27	38.6	11.8	672	6	CD884257 F1.116A21
C 28	38.6	11.8	701	6	CD875244 AZ03.104K
C 29	38.2	11.7	909	9	AG125251 Pan trogl
C 30	38.2	11.7	927	9	CL467547 SAIL_1271
C 31	38.2	11.7	1276	9	CL491442 SAIL_555
C 32	38	11.6	303	9	CL256246 FHCRG-GT-
C 33	38	11.6	684	6	CD446049 EL01T0206
C 34	38	11.6	896	6	CD439437 EL01N0524
C 35	38	11.6	1978	3	CNS0A7EC Arabidops
C 36	37.8	11.6	769	9	CNS032YU Arabidops
C 37	37.8	11.6	913	9	CL473483 SAIL_201
C 38	37.8	11.6	943	9	CL466335 SAIL_1254
C 39	37.8	11.6	1058	4	BG387202 602455932
C 40	37.8	11.6	1063	5	BX331486 BX331486
C 41	37.6	11.5	397	1	AJ437938 AJ437938
C 42	37.6	11.5	487	2	BE230675 99AS897 R
C 43	37.6	11.5	568	6	CD205709 HSI_15_G0
C 44	37.6	11.5	661	2	BE823637 GM700021A
C 45	37.6	11.5	724	6	CB627516 OS11EB021

## ALIGNMENTS

RESULT 1  
LOCUS CL467818/c 1086 bp DNA linear GSS 31-MAR-2004  
DEFINITION SAIL\_1276\_G09.v1 SAIL Collection Arabidopsis thaliana genomic clone  
SAIL\_1276\_G09.v1, genomic survey sequence.  
ACCESSION CL467818  
VERSION CL467818.1 GI:45870723  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1086)  
AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kinnerly,B., Mittel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.  
TITLE A high-throughput Arabidopsis reverse genetics system  
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)  
MEDLINE 22356987  
PUBMED 12468722  
COMMENT Contact: Sessions A  
Applied Trait Genetics  
Syngenta Biotechnology Inc.  
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
Email: allen.sessions@syngenta.com  
ABRC Stock Number CS847449; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).  
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .1086  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL\_1276\_G09.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

FEATURES  
source  
1. .1086  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL\_1276\_G09.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

## ORIGIN

Query Match 13.2%; Score 43.2; DB 9; Length 1086;  
Best Local Similarity 56.1%; Pred. No. 0.41;

Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 179 ACAGTCGCAAGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTCTGGGTCTTT 238  
 |||||  
 Db 570 AGAGTGGGAAGGTCCGAGGAGTAGAGTGTGGAAGGAGGAGGCGGCGGCGGCGGCGGT 511  
 |||||

Qy 239 TCTGCGGAGGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGT 298  
 |||||  
 Db 510 GGGGAGGAGGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 451  
 |||||

Qy 299 CGAGGTGAGGTGTGGCAGG 317  
 |||||

Db 450 GGAGGGAGGGGGGAGG 432  
 |||||

RESULT 2  
 CL496190/c  
 LOCUS  
 DEFINITION SAIL\_620\_G10.v3 1262 bp DNA linear GSS 01-APR-2004  
 SAIL\_620\_G10.v3, genomic survey sequence.

ACCESSION  
 VERSION  
 CL496190.1 GI:45988256

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE  
 1 (bases 1 to 1262)  
 AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,  
 Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D.,  
 Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B.,  
 Mittel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A.  
 A high-throughput Arabidopsis reverse genetics system  
 Plant Cell 14 (12), 2985-2994 (2002)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

COMMENT

Contact: Sessions A  
 Applied Trait Genetics  
 Syngenta Biotechnology Inc.  
 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
 Email: allen.sessions@syngenta.com  
 ABRC Stock Number C3826573; T-DNA left border flanking sequences of  
 Syngenta Arabidopsis Insertion Library (SAIL) lines are available  
 through the Arabidopsis Biological Resource Center (ABRC).  
 Sequences represent a pool of amplified genomic regions and not  
 single contiguous sequences.

Class: TDNA tagged.

Location/Qualifiers

1. .1262  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="SAIL\_620\_G10.v3"  
 /clone\_lib="SAIL Collection"  
 /note="T-DNA left border sequences were isolated using a  
 modified TAIL-PCR strategy"

ORIGIN

Query Match 13.0%; Score 42.6; DB 9; Length 1262;  
 Best Local Similarity 57.7%; Pred. No. 0.61;  
 Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 189 GGTAGCTGAGCACCGTGGCGGCGGCGAGCGGTGGCGGTGTTTCTGGCGGAG 248  
 |||||  
 Db 1217 GGGGGCGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1159  
 |||||

Qy 249 GTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGG 308  
 |||||  
 Db 1157 GTNTTGGTGGGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 1098  
 |||||

Qy 309 TGTGGCAGCG 318

Db 1097 TGGGCGCGGC 1088  
 |||||

RESULT 3

CC258768/c

LOCUS

DEFINITION CC258768 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.1 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258768

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-164A7"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EORI; Site 2: EORI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;  
 Best Local Similarity 59.2%; Pred. No. 1.1;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 204 GTGCGGCGCGACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 263  
 |||||  
 Db 741 GCGTGGTGGGTGGGTGGTGGCGGTGGGTGGTGGCGGTGGGTGGGTGGGTGGGTGG 682  
 |||||

Qy 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGGTTCGAGGTGAGGTGGCAGGCTTCAG 323  
 |||||

Db 681 GGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622  
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003  
 CH261-164A7\_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,  
 genomic survey sequence.

ACCESSION  
 VERSION  
 CC258769

KEYWORDS  
 GSS.

SOURCE  
 ORGANISM Gallus gallus (chicken)

Gallus gallus

REFERENCE

1 (bases 1 to 1164)

AUTHORS

Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT

Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TAGCACTCACTATAGGAGAG  
 Class: BAC ends  
 High quality sequence start: 186  
 High quality sequence stop: 475.

Location/Qualifiers

1. .1164  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:90



Db 247 GGGCTTGACCACTCGACAGAGCGGTTCGGAGGCGGTCTTGGGAAAGGCGACGCTTCG 188  
 QY 176 GTACAGTCGCAAGTAGCTAGACACCGTGGCGGCGGCGAGCGGTGGCGGTTCGGGGTT 235  
 Db 187 GACCTGGTGGCGGCGGTGCATCGCGGTGTAGAGCATGACAGGGCTCGGTGGGGAG 128  
 QY 236 GTTCTGGCGGAG 248  
 Db 127 GATGTTGGCGGAG 115

## RESULT 7

CA731965/c  
 LOCUS CA731965 632 bp mRNA linear EST 26-NOV-2002  
 DEFINITION wlpic.pk002.j18 wlpic Triticum aestivum cDNA clone wlpic.pk002.j18  
 5' end, mRNA sequence.

ACCESSION CA731965

VERSION CA731965

KEYWORDS CA731965.1 GI:25547563

SOURCE EST.

ORGANISM Triticum aestivum (bread wheat)

COMMENT Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

TITLE Poideae; Triticeae; Triticum.

JOURNAL 1 (bases 1 to 632)

COMMENT Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,

Miao, G., Caraher, N. and Hanafey, M.K.

Unpublished (2002)

Contact: Scott V. Tingey

Crop Genetics

E. I. DuPont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

Location/Qualifiers

1..632

/organism="Triticum aestivum"

/mol\_type="mRNA"

/db\_xref="taxon:4565"

/clone="wlpic.pk002.j18"

/tissue\_type="lemma and palea"

/lab\_host="DH10B"

/clone\_lib="wlpic"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:

XhoI; Wheat (Triticum aestivum, Hi Line) lemma and palea"

ORIGIN

Query Match

Best Local Similarity 12.3%; Score 40.2; DB 6; Length 632;

Mismatches 75; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 116 GGACCTGAGCGATCGGATCGACCGGATCGGAAACCTCTCGAGAAAGGCGGTCTAACCA 175

Db 546 GGGCTTGACCACTCGACAGAGCGGTTCGGAGGCGTCTTGGGAAAGGCGACGCTTCG 487

QY 176 GTACAGTCGCAAGTAGCTAGACACCGTGGCGGCGGCGAGCGGTTCGGGGTT 235

Db 486 GACCTGGTGGCGGCGGTGCATCGCGGTGTAGAGCATGACAGGGCTGCGGTTCGGGAG 427

QY 236 GTTCTGGCGGAG 248

Db 426 GATGTTGGCGGAG 414

RESULT 8

CA378248/c

LOCUS CA378248 598 bp mRNA linear EST 06-NOV-2002

DEFINITION 657011 NCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT42L01\_B\_F01 5',

mRNA sequence.

## ACCESSION

VERSION CA378248.1

KEYWORDS GI:24697932

SOURCE EST.

ORGANISM Oncorhynchus mykiss (rainbow trout)

COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 598)

REFERENCE Rexroad, C.E. 3rd, Lee, Y., Keele, J.W., Karamycheva, S., Brown, G.,

Koop, B., Gahr, S.A., Palti, Y. and Quackenbush, J.

Sequence analysis of a rainbow trout cDNA library and creation of a

gene index

Cytogenet. Genome Res. 102 (1-4), 347-354 (2003)

JOURNAL Contact: Rexroad CE

COMMENT USDA, ARS, National Center for Cool and Cold Water Aquaculture

11876 Lestown Road, Kearneysville, WV 25430, USA

Tel: 304 724 8340 x2129

Fax: 304 725 0351

Email: crexroad@nccwa.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim\_alt option. Vector identified by

cross match v0.990329.

Seq primer: AGCGGATACAAATTTCACACAGGA.

Location/Qualifiers

1..598

/organism="Oncorhynchus mykiss"

/mol\_type="mRNA"

/db\_xref="taxon:8022"

/clone="1RT42L01\_B\_F01"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="NCCWA 1RT"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from brain, gill, liver,

spleen, muscle, and kidney."

ORIGIN

Query Match

Best Local Similarity 12.2%; Score 40; DB 6; Length 598;

Mismatches 100; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 111 CCGAGGACCTGACGAGTCGGATCGACCGATCGGAAACCTCTCGAGAAAGCGCTCT 170

Db 308 CGGTAGAATATGATCGTGGAAACCTGGGGGCGAGCTCTCAAGCGGTGATGACGCTACG 249

QY 171 AACCACTGACAGTCGCAAGGTAGGTGACGACCGCTGGCGGCGGCGGCTGGCGTCG 230

Db 248 GACCGGTTTCCCTCGAAGGCTTTCCCGAGTGGGTGTCGATGTCAGCTTGACCTGGTTCG 189

QY 231 GGGTTGTTTCTGGCGGAGGTGCTCTGATGATGTAATTAAGTAGGGCGGTCTTGAGACGG 290

Db 188 CTGTCTCTGGCGAGCTTGTGTCGCCCGCTCGATGTCGATGAAGGAGGCGGTGAGGCGCAG 129

QY 291 CGGATGTCGAGGTGAGGTG 310

Db 128 GAGAAGCGGAGGCCAGATG 109

RESULT 9

CC286173/c

LOCUS CC286173 1177 bp DNA linear GSS 13-MAY-2003

DEFINITION CH261-29F4\_RM1.1 CH261 Gallus gallus genomic clone CH261-29F4,

genomic survey sequence.

ACCESSION CC286173

VERSION CC286173.1

KEYWORDS GI:30655433

SOURCE GSS.

ORGANISM Gallus gallus (chicken)

COMMENT Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 1177)

REFERENCE

**AUTHORS** Krenitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,  
 Warren,W., Graves,T., Mardis,E. and Wilson,R.  
**TITLE** Gallus gallus BAC End Reads  
**JOURNAL** Unpublished (2003)  
**COMMENT** Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TACGACTCACTATAGGGAGA  
 Class: BAC ends  
 High quality sequence start: 135  
 High quality sequence stop: 281.

**FEATURES**  
 Location/Qualifiers  
 1..1177  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-29F4"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /clone\_lib="CH261"  
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"

# ORIGIN

source

**Query Match** 12.2%; Score 40; DB 8; Length 1177;  
**Best Local Similarity** 55.9%; Pred. No. 3.2;  
**Matches** 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 189 GTTAGGCTGACACCGTGGCGGCGGAGCGGTGGCGGTGTTCTTCTGGCGGAG 248  
 |||||  
 DB 725 GCGAGGGGAGGGGGGTGGAGGGGGGAGGGGTGGAGGTGGGGGGGAGAGGGGAG 666  
 |||||

QY 249 GTGCTGCTGATGATGATTAATAAGTAGGCGGTCTTGAGACGCGGATGCTGAGGTGAGG 308  
 |||||  
 DB 665 GCGAGGGGTGGGATGGGGGTAAATGGGGGGGGGTAGAGAGGGCGGAGGTGGGGGAGAG 606  
 |||||

QY 309 TGTGCGAGCTTGAGA 324  
 |||||  
 DB 605 GCGGGGGGGGGGGGGA 590  
 |||||

**RESULT 10**  
**CL513605/c**  
**LOCUS** SAIL\_877\_H06.v1 936 bp DNA linear GSS 01-APR-2004  
**DEFINITION** SAIL\_877\_H06.v1 SAIL Collection Arabidopsis thaliana genomic clone  
**ACCESSION** CL513605.1 GI:46010925  
**VERSION** CL513605.1  
**KEYWORDS** GSS.

**ORGANISM** Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 936)  
**REFERENCE** Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,  
 Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,  
 Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,  
 Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.  
 A high-throughput Arabidopsis reverse genetics system  
**AUTHORS** Plant Cell 14 (12), 2985-2994 (2002)  
 22356987  
 12468722

**TITLE** Sessions A  
**JOURNAL** Applied Trait Genetics  
**MEDLINE** Syngenta Biotechnology Inc.  
**PUBMED** 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
**COMMENT** Email: allen.sessions@syngenta.com  
 ABRC Stock Number CS839484; T-DNA left border flanking sequences of

Syngenta Arabidopsis Insertion Library (SAIL) lines are available  
 through the Arabidopsis Biological Resource Center (ABRC).  
 Sequences represent a pool of amplified genomic regions and not  
 single contiguous sequences.  
 Class: T-DNA tagged

# FEATURES

source

Location/Qualifiers  
 1..936  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="SAIL\_877\_H06.v1"  
 /clone\_lib="SAIL Collection"  
 /note="T-DNA left border sequences were isolated using a  
 modified TAIL-PCR strategy"

# ORIGIN

**Query Match** 12.1%; Score 39.6; DB 9; Length 936;  
**Best Local Similarity** 56.0%; Pred. No. 4;  
**Matches** 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 188 AGGTAGGCTGACACCGTGGCGGCGGAGCGGTGGCGGTGTTCTTCTGGCGGA 247  
 |||||  
 DB 875 AGATGGGGGAGCCCGAGGGGGGGGAGCGGGGGGAGGTGGGGGGGGGGGG 816  
 |||||

QY 248 GGTCTGCTGATGATGATTAATAAGTAGCGCGTCTTGAGACGCGGATGCTCGAGGTGAG 307  
 |||||  
 DB 815 GGGGGAGATGCTGTAGGCAAGAGATGGAGCAGAGGGGTGGGTAGGGGGGGGG 756  
 |||||

QY 308 GTGTGGCAGGCTTG 321  
 |||||  
 DB 755 GGGGGGAGGGGTGG 742  
 |||||

# RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Helianthus annuus (common sunflower)

Helianthus annuus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
 Heliantheae; Helianthus.  
 1 (bases 1 to 511)  
**REFERENCE** Tamboirindeguy,C., Ben,C., Liboz,T. and Gentzbbittel,L.  
**AUTHORS** Sequence evaluation of four specific cDNA libraries for  
**TITLE** developmental genomics of sunflower  
**JOURNAL** Mol. Genet. Genomics 271 (3), 367-375 (2004)  
**COMMENT** Contact: Gentzbbittel L  
 Laboratoire de Biotechnologie et Amélioration des Plantes  
 Institut National Polytechnique de Toulouse - Ecole National  
 Supérieure Agronomique de Toulouse  
 IPR40; Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,  
 Auzeville, CASTANET TOLOSAN 31326, France.

# FEATURES

source

Location/Qualifiers  
 1..511  
 /organism="Helianthus annuus"  
 /mol\_type="mRNA"  
 /cultiivar="Emil"  
 /db\_xref="taxon:4232"  
 /clone="HapIR205H12"  
 /tissue type="hypocotyls"  
 /cell type="protoplasts"  
 /dev\_stage="1-5 days"  
 /clone\_lib="HapIR2"

# ORIGIN

**Query Match** 12.0%; Score 39.2; DB 1; Length 511;







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 9289.17 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-32  
Perfect score: 1240  
Sequence: 1 ggatccactctcttcgcgcacgttcacagtcgcaagatctc 1240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_srs.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1240	6	BD268221 Adenoviru
2	1239	99.9	7231	6	BD268252 Adenoviru
3	1239	99.9	7960	6	BD268233 Adenoviru
4	1239	98.9	7989	6	BD268236 Adenoviru
5	1239	99.9	8383	6	BD268232 Adenoviru
6	1239	99.9	8484	6	BD268253 Adenoviru
7	1135	91.5	10332	6	A83180 Sequence 13
8	1135	91.5	10332	6	BD082846 Method an
9	1135	91.5	11570	14	AD5001
10	1135	91.5	31976	6	CQ854904 Sequence
11	1135	91.5	31976	6	CQ854905 Sequence
12	1135	91.5	32480	6	BD268216 Adenoviru
13	1135	91.5	32798	6	AX343138 Sequence
14	1135	91.5	32798	6	AX382187 Sequence
15	1135	91.5	32802	6	CQ854906 Sequence
16	1135	91.5	33007	12	AF323988 EGFP expr
17	1135	91.5	33476	12	AY370909 Expressio
18	1135	91.5	33592	6	AX084504 Sequence
19	1135	91.5	33699	6	AX084506 Sequence

20	1135	91.5	33988	6	AX084517 Sequence
21	1135	91.5	34303	6	AR091536 Sequence
22	1135	91.5	34303	6	AR102229 Sequence
23	1135	91.5	34303	6	AR230727 Sequence
24	1135	91.5	34341	6	AX084505 Sequence
25	1135	91.5	34448	6	AX084507 Sequence
26	1135	91.5	34737	6	AX084518 Sequence
27	1135	91.5	35408	6	AR163568 Sequence
28	1135	91.5	35408	6	AR166442 Sequence
29	1135	91.5	35724	6	AX084516 Sequence
30	1135	91.5	35764	12	AY046510 Adenovira
31	1135	91.5	35871	6	AR403724 Sequence
32	1135	91.5	35934	14	AY339865 Human ade
33	1135	91.5	35935	6	AR091533 Sequence
34	1135	91.5	35935	6	AR102226 Sequence
35	1135	91.5	35935	6	AR116313 Sequence
36	1135	91.5	35935	6	CQ854907 Sequence
37	1135	91.5	35935	6	AR230724 Sequence
38	1135	91.5	35935	6	AX451988 Sequence
39	1135	91.5	35935	6	AX683770 Sequence
40	1135	91.5	35935	14	ADRCOMPGEN M73260 Mastadenovi
41	1135	91.5	35978	6	AR403723 Sequence
42	1135	91.5	36114	6	AX084519 Sequence
43	1135	91.5	36154	6	AX468857 Sequence
44	1135	91.5	36154	6	AX468865 Sequence
45	1135	91.5	36620	6	AR534337 Sequence

ALIGNMENTS

RESULT 1	BD268221	Adenovirus vector, packaging cell line, composition and method for production and use.	1240 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD268221					
DEFINITION	ADENOVIRUS	Adenovirus vector, packaging cell line, composition and method for production and use.				
ACCESSION	BD268221.1	GI:33077989				
VERSION	JP 2002534130-A/25					
KEYWORDS	unidentified adenovirus					
SOURCE	unidentified adenovirus					
ORGANISM	Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.					
REFERENCE	1 (bases 1 to 1240)					
AUTHORS	Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.					
TITLE	Adenovirus vector, packaging cell line, composition and method for production and use					
JOURNAL	Patent: JP 2002534130-A 25 15-OCT-2002; NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE					
COMMENT	OS Adenovirus					
	PN JP 2002534130-A/25					
	PD 15-OCT-2002					
	PF 14-JAN-2000 JP 2000593765					
	PR 14-JAN-1999 US 60/115920					
	PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCCHENKO					
	PC C12N15/09,A61K35/76,A61P35/00,A61P43/00,A61P43/00, A61P43/00, C12N5/10,					
	PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00, C12N5/00 CC					
	Adenovirus vector, packaging cell line, composition and method for					
	CC production and use					
	CC Key					
	CC Location/Qualifiers					
	FT source					
	FT 1. .1240					
	FT /organism='Adenovirus'.					
	FT Location/Qualifiers					
	1. .1240					
	/organism='unidentified adenovirus'					
	/mol_type='genomic DNA'					
	/db_xref='taxon:10535'					
FEATURES	source					
	1. .1240					
	/organism='unidentified adenovirus'					
	/mol_type='genomic DNA'					
	/db_xref='taxon:10535'					
ORIGIN						

Query Match	100.0%;	Score 1240;	DB 6;	Length 1240;	
Best Local Similarity	100.0%;	Pred. No. 1e-237;			
Matches 1240;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	GGATCCACTCTCTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT	60		
Db	1	GGATCCACTCTCTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT	60		
Qy	61	CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTGAGGGTGCGCGATCCATCTGGTCAAGAAA	120		
Db	61	CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTGAGGGTGCGCGATCCATCTGGTCAAGAAA	120		
Qy	121	GATATTACCTGGCGCGCGGTGATCCCTTTGAGGGTGCGCGATCCATCTGGTCAAGAAA	180		
Db	121	GATATTACCTGGCGCGCGGTGATCCCTTTGAGGGTGCGCGATCCATCTGGTCAAGAAA	180		
Qy	181	GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCGGTAGAGGGCGTTGGACAGCAA	240		
Db	181	GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCGGTAGAGGGCGTTGGACAGCAA	240		
Qy	241	CTTGGCGATGAGCGCGAGGTTGGTTTGTGCGGATCGCGCGCTCTTTGGCGCGCAT	300		
Db	241	CTTGGCGATGAGCGCGAGGTTGGTTTGTGCGGATCGCGCGCTCTTTGGCGCGCAT	300		
Qy	301	GTTTACCTGCGAGTATTCGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGCGGCTC	360		
Db	301	GTTTACCTGCGAGTATTCGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGCGGCTC	360		
Qy	361	GTGCGGCGACAGGTGACGCGCGCAACCGCGTTGTGCGAGGGTGACAAGGTCAACGCTCGT	420		
Db	361	GTGCGGCGACAGGTGACGCGCGCAACCGCGTTGTGCGAGGGTGACAAGGTCAACGCTCGT	420		
Qy	421	GGCTACCTCTCCGCTAGCGCTGTTGGTTCAGCAGAGGCGCGCGCTTTGCGCGAGCA	480		
Db	421	GGCTACCTCTCCGCTAGCGCTGTTGGTTCAGCAGAGGCGCGCGCTTTGCGCGAGCA	480		
Qy	481	GAATCGCGGTAGGGGTCTAGCTGCGTCTCGTCCGGGGGCTGCGTCCACGGTAAAGAC	540		
Db	481	GAATCGCGGTAGGGGTCTAGCTGCGTCTCGTCCGGGGGCTGCGTCCACGGTAAAGAC	540		
Qy	541	CCCGGCGACAGCGCGCGTCAAGTAGTCTATCTTGATCTCTTCAAGTCTAGCGCCTG	600		
Db	541	CCCGGCGACAGCGCGCGTCAAGTAGTCTATCTTGATCTCTTCAAGTCTAGCGCCTG	600		
Qy	601	CTGCCATCGCGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGACCCCATGGCAT	660		
Db	601	CTGCCATCGCGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGACCCCATGGCAT	660		
Qy	661	GGGTGGGTGAGCGGAGGCGTACATGCCGAATGTCTGTAACGTAGAGGGGCTCTCT	720		
Db	661	GGGTGGGTGAGCGGAGGCGTACATGCCGAATGTCTGTAACGTAGAGGGGCTCTCT	720		
Qy	721	GAGTATTCGAAGATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC	780		
Db	721	GAGTATTCGAAGATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC	780		
Qy	781	GTAATAGTCTGCGAGGAGCGAGAGGTGCGGAACGAGTTGCTACGGGCGGCTGCTC	840		
Db	781	GTAATAGTCTGCGAGGAGCGAGAGGTGCGGAACGAGTTGCTACGGGCGGCTGCTC	840		
Qy	841	TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG	900		
Db	841	TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG	900		
Qy	901	GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCAACGAGGAGGCGGTAGGA	960		
Db	901	GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCAACGAGGAGGCGGTAGGA	960		
Qy	961	GTGCGGAGCTTTGTAACAGCTCGCGGTGACCTGACAGTCTAGGGCGCAGTACTCCAG	1020		
Db	961	GTGCGGAGCTTTGTAACAGCTCGCGGTGACCTGACAGTCTAGGGCGCAGTACTCCAG	1020		
Qy	1021	GGTTTCTTGATGATGCATCTTATCTCTGCTCCCTTTTTTTTCCACAGCTCGCGGTTGAG	1080		

Db	1021	GGTTTCTTGATGATGCATCTTATCTCTGCTCCCTTTTTTTTCCACAGCTCGCGTTGAG	1080		
Qy	1081	GACAACTCTTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG	1140		
Db	1081	GACAACTCTTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG	1140		
Qy	1141	AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCCACCGATCGGAAACCC	1200		
Db	1141	AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCCACCGATCGGAAACCC	1200		
Qy	1201	TCTCGAAGAGGCGCTTAAACAGTCACAGTCGCAAGATCT	1240		
Db	1201	TCTCGAAGAGGCGCTTAAACAGTCACAGTCGCAAGATCT	1240		
RESULT 2					
BD268252		7231 bp	DNA	linear	PAT 17-JUL-2003
LOCUS					
DEFINITION					Adenovirus vector, packaging cell line, composition and method for production and use.
ACCESSION					BD268252
VERSION					BD268252.1
KEYWORDS					JP 2002534130-A/56.
SOURCE					synthetic construct
ORGANISM					synthetic construct
REFERENCE					1 (bases 1 to 7231)
AUTHORS					Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.
TITLE					Adenovirus vector, packaging cell line, composition and method for production and use
JOURNAL					Patent: JP 2002534130-A 56 15-OCT-2002;
COMMENT					NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
					OS Artificial Sequence
					PN JP 2002534130-A/56
					PD 15-OCT-2002
					PF 14-JAN-2000 JP 2000593765
					PI 14-JAN-1999 US 60/115920
					PR GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCHENKO
					PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
					PC C12N5/10,
					PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
					Description of Artificial Sequence: plasmid
					FT Key Location/Qualifiers
					FT source 1..7231 /organism='Artificial Sequence'.
FEATURES					
source					1..7231
					/organism="synthetic construct"
					/mol_type="genomic DNA"
					/db_xref="taxon:32630"
ORIGIN					
Query Match					99.9%; Score 1239; DB 6; Length 7231;
Best Local Similarity					100.0%; Pred. No. 1.4e-237;
Matches 1239;					Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT	60		
Db	849	GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT	908		
Qy	61	CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAACGAGGAGATT	120		
Db	909	CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAACGAGGAGATT	968		
Qy	121	GATATTACCTGGCGCGCGGTGATCCCTTTGAGGGTGCGCGATCCATCTGGTCAAGAAA	180		
Db	969	GATATTACCTGGCGCGCGGTGATCCCTTTGAGGGTGCGCGATCCATCTGGTCAAGAAA	1028		
Qy	181	GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCGGTAGAGGGCGTTGGACAGCAA	240		

Db 1029 GACAACTTTTGTGTCAGCTTGTGTCGCAACCGATAGAGGGCGTTGACAGCA 1088  
Qy 241 CTTGCGATGAGCGCAGCGTTTGGTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 300  
Db 1089 CTTGCGATGAGCGCAGCGTTTGGTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 1148  
Qy 301 GTTTAGCTGACGATTTCCGCGGCAACCGACCGCATTCGGGAAAGACGGTGGTGGCTC 360  
Db 1149 GTTTAGCTGACGATTTCCGCGGCAACCGACCGCATTCGGGAAAGACGGTGGTGGCTC 1208  
Qy 361 GTGCGGCAACCGATTCGCGGCAACCGCGGTTGTGCGAGGTGACAAAGTCAACGCTGT 420  
Db 1209 GTGCGGCAACCGATTCGCGGCAACCGCGGTTGTGCGAGGTGACAAAGTCAACGCTGT 1268  
Qy 421 GGCTACCTCTCCGCGTAGCGCTGCTGTGTCAGCAGAGCGCGCCCTTCCGCGAGCA 480  
Db 1269 GGCTACCTCTCCGCGTAGCGCTGCTGTGTCAGCAGAGCGCGCCCTTCCGCGAGCA 1328  
Qy 481 GAATGGCGTAGGGGTCTAGCTGCGTCTGCTGTCGGGGGGTCTGCTCCACGTTAAAGAC 540  
Db 1329 GAATGGCGTAGGGGTCTAGCTGCGTCTGCTGTCGGGGGGTCTGCTCCACGTTAAAGAC 1388  
Qy 541 CCGGCGCAGCGCGCGTCAAGTGTCTATCTTGCATCTCTCAAGTCTAGCGCTG 600  
Db 1389 CCGGCGCAGCGCGCGTCAAGTGTCTATCTTGCATCTCTCAAGTCTAGCGCTG 1448  
Qy 601 CTGCCATGCGCGCGCGCGTCAAGTGTCTATCTTGCATCTCTCAAGTCTAGCGCTG 720  
Db 1449 CTGCCATGCGCGCGCGCGTCAAGTGTCTATCTTGCATCTCTCAAGTCTAGCGCTG 1508  
Qy 661 GGGTGGGTGAGCGCGGAGCGTCAATGCCGCAATGTCTGTAACGTAAGGGGTCTCT 780  
Db 1509 GGGTGGGTGAGCGCGGAGCGTCAATGCCGCAATGTCTGTAACGTAAGGGGTCTCT 1568  
Qy 721 GAGTATCCAAATATGATAGGTAGCTATCTTCCACCGCGATGCTGGCGCGCACGTAATC 1628  
Db 1569 GAGTATCCAAATATGATAGGTAGCTATCTTCCACCGCGATGCTGGCGCGCACGTAATC 1688  
Qy 781 GTATAGTTCTGCGGAGGAGCGAGAGGTGCGGACCGAGTTGCTACGCGCGGCTGCTC 840  
Db 1629 GTATAGTTCTGCGGAGGAGCGAGAGGTGCGGACCGAGTTGCTACGCGCGGCTGCTC 1688  
Qy 841 TGCTCGAAGACTATCTGCTCAAGATGGCATGTGAGTTGGATGATATGTTTGGACGCTG 900  
Db 1689 TGCTCGAAGACTATCTGCTCAAGATGGCATGTGAGTTGGATGATATGTTTGGACGCTG 1748  
Qy 901 GAAGAGTTGAAGCTGGGCTGTGAGACCTACCGGTGACGACGAGAGGCGGTAGGA 960  
Db 1749 GAAGAGTTGAAGCTGGGCTGTGAGACCTACCGGTGACGACGAGAGGCGGTAGGA 1808  
Qy 961 GTGCGCAGCTTTGACAGCTCGCGGTGACCTGACGTTCTAGGGCGCAGTAGTCCAG 1020  
Db 1809 GTGCGCAGCTTTGACAGCTCGCGGTGACCTGACGTTCTAGGGCGCAGTAGTCCAG 1868  
Qy 1021 GGTTCCTTGATGATGATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1080  
Db 1869 GGTTCCTTGATGATGATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1928  
Qy 1081 GACAACTCTTCGCGGTCTTTCCAGTACTCTGTTGATCGGAAACCGCTCGGCTCCGACG 1140  
Db 1929 GACAACTCTTCGCGGTCTTTCCAGTACTCTGTTGATCGGAAACCGCTCGGCTCCGACG 1988  
Qy 1141 AGATCGTACTCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAAACCC 1200  
Db 1989 AGATCGTACTCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAAACCC 2048  
Qy 1201 TCTCGAAGAGGCGTCTAACAGTACAGTTCGCAAGATC 1239  
Db 2049 TCTCGAAGAGGCGTCTAACAGTACAGTTCGCAAGATC 2087

LOCUS BD268233 7960 bp DNA linear PAT 17-JUL-2003  
DEFINITION Adenovirus vector, packaging cell line, composition and method for  
production and use.  
ACCESSION BD268233  
VERSION BD268233.1 GI:33078001  
KEYWORDS JP 2002534130-A/37.  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 7960)  
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and  
Skripchenko,Y.  
TITLE Adenovirus vector, packaging cell line, composition and method for  
production and use  
JOURNAL Patent: JP 2002534130-A 37 15-OCT-2002;  
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE  
COMMENT OS Artificial Sequence  
PN JP 2002534130-A/37  
PD 15-OCT-2002 JP 2000593765  
PF 14-JAN-2000 JP 2000593765  
PR 14-JAN-1999 US 60/115920  
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI  
SUSAN C STEVENSON,YELENA SKRIPCENKO  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,  
PC C12N5/10,  
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC  
Description of Artificial Sequence: plasmid  
FH Key Location/Qualifiers  
FT source 1..7960  
/organism='Artificial Sequence'.  
FEATURES  
source  
1..7960  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
ORIGIN  
Query Match 99.9%; Score 1239; DB 6; Length 7960;  
Best Local Similarity 100.0%; Pred.No.1.4e-237;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGATCCACTCTCTCCGCATCGCTCTGCGAGGCGCAGCTGTGGGGTGAATCTCCCT 60  
Db 929 GGATCCACTCTCTCCGCATCGCTCTGCGAGGCGCAGCTGTGGGGTGAATCTCCCT 988  
Qy 61 CTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGAGGATTT 120  
Db 989 CTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGAGGATTT 1048  
Qy 121 GATATTCACTTGGCCCGCGGTGATGCCCTTTCAGGGTGGCCGCATCATCTGTCAGAAA 180  
Db 1049 GATATTCACTTGGCCCGCGGTGATGCCCTTTCAGGGTGGCCGCATCATCTGTCAGAAA 1108  
Qy 181 GACAACTTTTGTGTCAGCTTGGTGGCAAAACGACCGGTAGAGGGGTGGAGACGAA 240  
Db 1109 GACAACTTTTGTGTCAGCTTGGTGGCAAAACGACCGGTAGAGGGGTGGAGACGAA 1168  
Qy 241 CTTGCGATGAGCGCAGGTTTGTGTCGCGATCGCGCGCTCTCTTGGCCCGCAT 300  
Db 1169 CTTGCGATGAGCGCAGGTTTGTGTCGCGATCGCGCGCTCTCTTGGCCCGCAT 1228  
Qy 301 GTTTAGCTGACGATTTCCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 360  
Db 1229 GTTTAGCTGACGATTTCCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 1288  
Qy 361 GTGCGGCAACGAGGTGACGCGCAACCGCGGTTGTGCAAGGTGACAAAGTCAACGCTGT 420  
Db 1289 GTGCGGCAACGAGGTGACGCGCAACCGCGGTTGTGCAAGGTGACAAAGTCAACGCTGT 1348  
Qy 421 GGCTACCTCTCGCGTAGCGCTCGTTGTCAGCAGAGGCGCGCTCTTGGCGGAGCA 480  
Db 1349 GGCTACCTCTCGCGTAGCGCTCGTTGTCAGCAGAGGCGCGCTCTTGGCGGAGCA 1408

```
QY 481 GAATGCGGTAGGGGTCTAGCTGCTCTCGTCCGGGGTCTGCTCCACGGTAAGAC 540
Db 1409 GAATGCGGTAGGGGTCTAGCTGCTCTCGTCCGGGGTCTGCTCCACGGTAAGAC 1468
QY 541 CCCGGGACGAGCGCGGTGAAAGTAGTCTATCTTTGTCATCTTCCAAAGTCTAGCGCTG 600
Db 1469 CCCGGGACGAGCGCGGTGAAAGTAGTCTATCTTTGTCATCTTCCAAAGTCTAGCGCTG 1528
QY 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATCGGTGAGTGGGGACCCCATGGCAT 660
Db 1529 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATCGGTGAGTGGGGACCCCATGGCAT 1588
QY 661 GGGGTGGGTGACGCGGGAGGCTACATGCCCAATGTCGTAAGTAGAGGGGTCTCT 720
Db 1589 GGGGTGGGTGACGCGGGAGGCTACATGCCCAATGTCGTAAGTAGAGGGGTCTCT 1648
QY 721 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
Db 1649 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1708
QY 781 GTATAGTTCTGCGAGGAGCGAGAGGTGCGGACCGAGTTGCTTACGGGGGGTGTCTC 840
Db 1709 GTATAGTTCTGCGAGGAGCGAGAGGTGCGGACCGAGTTGCTTACGGGGGGTGTCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTCAAGATGGCATGGTGGATGATATGTTGGACGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTCAAGATGGCATGGTGGATGATATGTTGGACGCTG 1828
QY 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCACGCACGAAGGAGCGTAGGA 960
Db 1829 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCACGCACGAAGGAGCGTAGGA 1888
QY 961 GTGCGGAGCTTTGTGAACAGCTCGGCGGTGACCTGACGCTAGGGGCGCAGTAGTCCAG 1020
Db 1889 GTGCGGAGCTTTGTGAACAGCTCGGCGGTGACCTGACGCTAGGGGCGCAGTAGTCCAG 1948
QY 1021 GGTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACACAGCTCGCGTTGAG 1080
Db 1949 GGTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACACAGCTCGCGTTGAG 2008
QY 1081 GACAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 1140
Db 2009 GACAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 2068
QY 1141 AGATCGTACTCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACGATCGGAACACC 1200
Db 2069 AGATCGTACTCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACGATCGGAACACC 2128
QY 1201 TCTCGAGAAAGGCTTAACCAAGTCACAGTCGCAAGATC 1239
Db 2129 TCTCGAGAAAGGCTTAACCAAGTCACAGTCGCAAGATC 2167
```

```
RESULT 4
LOCUS BD268236 7989 bp DNA linear PAT 17-JUL-2003
DEFINITION Adenovirus vector, packaging cell line, composition and method for
production and use.
ACCESSION BD268236
VERSION BD268236.1 GI:33078004
KEYWORDS JP 2002534130-A/40
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
AUTHORS 1 (bases 1 to 7989)
Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
Skrupchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for
production and use
JOURNAL. Patent: JP 2002534130-A 40 15-OCT-2002;
NOVARTIS AG,THE SCRIPHS RESEARCH INSTITUTE
COMMENT OS Artificial Sequence
PN JP 2002534130-A/40
```

```
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PI 14-JAN-1999 US 60/115920
GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
C12N5/10,
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FT Key Location/Qualifiers
FT source 1..7989
/organism='Artificial Sequence'
FEATURES
source
1..7989
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 99.9%; Score 1239; DB 6; Length 7989;
Best Local Similarity 100.0%; Pred. No. 1.4e-237;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTCTTCCGATCGCTGTCTGCGAGGCGCAGCTGTTGGGTGAGTACTCCCT 60
Db 929 GGATCCACTCTCTTCCGATCGCTGTCTGCGAGGCGCAGCTGTTGGGTGAGTACTCCCT 988
QY 61 CTGAAAACGCGGCATGACTTCTGCGCTAAGATTGCTCAGTTTCCAAAACGAGGAGGATT 120
Db 989 CTGAAAACGCGGCATGACTTCTGCGCTAAGATTGCTCAGTTTCCAAAACGAGGAGGATT 1048
QY 121 GATATTACCTGGGCGCGGTGATGCTTTCAGGGTGGCGGCATCCATCTGCTCAGAAA 180
Db 1049 GATATTACCTGGGCGCGGTGATGCTTTCAGGGTGGCGGCATCCATCTGCTCAGAAA 1108
QY 181 GACAACTCTTTTGTGTGCAAGCTTGTGGCAACGACCCGTAGAGGGGTGGGACAGCAA 240
Db 1109 GACAACTCTTTTGTGTGCAAGCTTGTGGCAACGACCCGTAGAGGGGTGGGACAGCAA 1168
QY 241 CTTGCGCATGAGCGCAGGGTTTGGTTTGTGTCGATCGCGCGCTCTTTGGCGCGCAT 300
Db 1169 CTTGCGCATGAGCGCAGGGTTTGGTTTGTGTCGATCGCGCGCTCTTTGGCGCGCAT 1228
QY 301 GTTTAGTGCACGTAATTGCGCGCGCAACGACCCGCAATTCGGGAAAGACGGTGGCGCTC 360
Db 1229 GTTTAGTGCACGTAATTGCGCGCGCAACGACCCGCAATTCGGGAAAGACGGTGGCGCTC 1288
QY 361 GTGCGGCAACAGGTGCAACGCGCTTGTGAGGGTGCACAGGTCAACGCTGGT 420
Db 1289 GTGCGGCAACAGGTGCAACGCGCTTGTGAGGGTGCACAGGTCAACGCTGGT 1348
QY 421 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGCGCGCGCTTTGGCGAGCA 480
Db 1349 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGCGCGCGCTTTGGCGAGCA 1408
QY 481 GAATGGCGGTAGGGGTCTAGCTCGCTCTGCTCGGGGGGTCTGGGTCCAGGTAAAGAC 540
Db 1409 GAATGGCGGTAGGGGTCTAGCTCGCTCTGCTCGGGGGGTCTGGGTCCAGGTAAAGAC 1468
QY 541 CCCGGGACGAGCGCGCTCGTCAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCCTG 600
Db 1469 CCCGGGACGAGCGCGCTCGTCAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCCTG 1528
QY 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGACCCCATGGCAT 660
Db 1529 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGACCCCATGGCAT 1588
QY 661 GGGGTGGGTGAGCGGGAGGCTACATGCCCAATGTCGTAAGTAGAGGGGTCTCT 720
Db 1589 GGGGTGGGTGAGCGGGAGGCTACATGCCCAATGTCGTAAGTAGAGGGGTCTCT 1648
QY 721 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
Db 1649 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1708
```

Db 1649 GAGTATTCAGAGATATGTAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGTAATC 1708  
Qy 781 GTATAGTTCTGTCGAGGAGAGCGAGAGTTCGGGACCGAGTTGTCTACGGCGGGCTGCTC 840  
Db 1709 GTATAGTTCTGTCGAGGAGAGCGAGAGTTCGGGACCGAGTTGTCTACGGCGGGCTGCTC 1768  
Qy 841 TCTCTGGAAGACTATCTGCTCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTG 900  
Db 1769 TCTCTGGAAGACTATCTGCTCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTG 1828  
Qy 901 GAAGACGTTGAAGCTGGCTGTGTGAGACCTACCGGTCACCGACGAGGAGGCTGAGGA 960  
Db 1829 GAAGACGTTGAAGCTGGCTGTGTGAGACCTACCGGTCACCGACGAGGAGGCTGAGGA 1888  
Qy 961 GTCTGGCGAGCTTGTGACAGCTTCGGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1020  
Db 1889 GTCTGGCGAGCTTGTGACAGCTTCGGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1948  
Qy 1021 GGTTCCTTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1080  
Db 1949 GGTTCCTTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 2008  
Qy 1081 GACAACTCTTCGGCGTCTTTCAGTACTCTTGGATCGGAACCGTTCGGCTTCGGAACG 1140  
Db 2009 GACAACTCTTCGGCGTCTTTCAGTACTCTTGGATCGGAACCGTTCGGCTTCGGAACG 2068  
Qy 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTTCGCACTCGACCGGATCGGAACCG 1200  
Db 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTTCGCACTCGACCGGATCGGAACCG 2128  
Qy 1201 TCTCGAAGAGCGCTTACCAAGTACAGTTCGCAAGATC 1239  
Db 2129 TCTCGAAGAGCGCTTACCAAGTACAGTTCGCAAGATC 2167

RESULT 5  
BD268232  
LOCUS Adenovirus vector, packaging cell line, composition and method for production and use. 8383 bp DNA linear PAT 17-JUL-2003  
DEFINITION  
ACCESSION BD268232  
VERSION BD268232.1 GI:33078000  
KEYWORDS JP 2002534130-A/36.  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 8383)  
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.  
TITLE Adenovirus vector, packaging cell line, composition and method for production and use.  
JOURNAL Patent: JP 2002534130-A 36 15-OCT-2002;  
COMMENT NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE  
OS Artificial Sequence  
PN JP 2002534130-A/36  
PD 15-OCT-2002  
PF 14-JAN-2000 JP 2000593765  
PR 14-JAN-1999 US 60/115920  
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI  
PC SUSAN C STEVENSON,YELENA SKRIPCHENKO  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,  
PC C12N5/10.  
PC C12N7/00,C12Q1/68,C01N33/53,C01N33/566,C12N15/00,C12N5/00 CC  
Description of Artificial Sequence: plasmid  
FH Key Location/Qualifiers  
FT source 1..8383  
FT Location/Qualifiers  
FEATURES  
source 1..8383  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
ORIGIN

Query Match 99.9%; Score 1239; DB 6; Length 8383;  
Best Local Similarity 100.0%; Pred. No. 1.4e-237;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGATCCACTCTCTCCGCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 60  
Db 907 GGATCCACTCTCTCCGCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 966  
Qy 61 CTGAAAACGGCGCATGACTTCTGCGCTAAGATTTGTCAGTTTCCAAAAACGAGGAGGATTT 120  
Db 967 CTGAAAACGGCGCATGACTTCTGCGCTAAGATTTGTCAGTTTCCAAAAACGAGGAGGATTT 1026  
Qy 121 GATATTACCTTGGCCCGCGGTGATGCTTTGAGGGTGGCCGCATCCATCTGCTCAGAAAA 180  
Db 1027 GATATTACCTTGGCCCGCGGTGATGCTTTGAGGGTGGCCGCATCCATCTGCTCAGAAAA 1086  
Qy 181 GACATCTTTTGTGTCAGCTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAA 240  
Db 1087 GACATCTTTTGTGTCAGCTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAA 1146  
Qy 241 CTTCGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 300  
Db 1147 CTTCGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 1206  
Qy 301 GTTTAGCTGCACTATTCGCGCGCAACGCAACCGCCATTCGGGAAAGACGGTGGTGGCTC 360  
Db 1207 GTTTAGCTGCACTATTCGCGCGCAACGCAACCGCCATTCGGGAAAGACGGTGGTGGCTC 1266  
Qy 361 GTTCGGGACAGAGTGCACGCGCAACCGCGGTGTGTGAGGGTGAACAAGTCAACGCTGGT 420  
Db 1267 GTTCGGGACAGAGTGCACGCGCAACCGCGGTGTGTGAGGGTGAACAAGTCAACGCTGGT 1326  
Qy 421 GGTACCTCTCCGCGTAGGCGCTGTTGTCAGAGAGGCGCGCCCTTGCAGGAGCA 480  
Db 1327 GGTACCTCTCCGCGTAGGCGCTGTTGTCAGAGAGGCGCGCCCTTGCAGGAGCA 1386  
Qy 481 GAATGGCGGTAGGGGCTTAGCTGCGTCTCGTCCGGGGGCTGTCGTCACCGGTAAAGAC 540  
Db 1387 GAATGGCGGTAGGGGCTTAGCTGCGTCTCGTCCGGGGGCTGTCGTCACCGGTAAAGAC 1446  
Qy 541 CCCGGGACAGAGCGCGCTGCAAGTAGTCTATCTTTGTCATCTTGCATCTCTGCAAGTCTAGCGCTG 600  
Db 1447 CCCGGGACAGAGCGCGCTGCAAGTAGTCTATCTTTGTCATCTTGCATCTCTGCAAGTCTAGCGCTG 1506  
Qy 601 CTGCGCATCGCGCGCGCAAGCGCGCTGATGAGTTGAGTGGGGACCCCATGGCAT 660  
Db 1507 CTGCGCATCGCGCGCGCAAGCGCGCTGATGAGTTGAGTGGGGACCCCATGGCAT 1566  
Qy 661 GGGGTGGGTGAGCGCGAGGCGTACATCCCGCAATGTCTGTAACGTAAGAGGGGCTCTCT 720  
Db 1567 GGGGTGGGTGAGCGCGAGGCGTACATCCCGCAATGTCTGTAACGTAAGAGGGGCTCTCT 1626  
Qy 721 GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGGATGTGGCGCGCATTAATC 780  
Db 1627 GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGGATGTGGCGCGCATTAATC 1686  
Qy 781 GTATAGTTCTGCGAGGAGCGAGGAGTCCGACCGAGGTTGCTACCGGCGGGCTGCTC 840  
Db 1687 GTATAGTTCTGCGAGGAGCGAGGAGTCCGACCGAGGTTGCTACCGGCGGGCTGCTC 1746  
Qy 841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTTGAACGCTG 900  
Db 1747 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTTGAACGCTG 1806  
Qy 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCGACGAGAGGGCTAGGA 960  
Db 1807 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCGAGAGGGCTAGGA 1866  
Qy 961 GTCGCGAGCTTGTGACAGCTCGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAG 1020  
Db 1867 GTCGCGAGCTTGTGACAGCTCGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAG 1926

QY 1021 GGTTCCTTGATGATCATCTTATCTGTCCTCTTTTTCACAGCTCGCGTTGAG 1080  
Db |||||||  
1927 GGTTCCTTGATGATCATCTTATCTGTCCTCTTTTTCACAGCTCGCGTTGAG 1986  
QY 1081 GACAACTCTTTCGGCGTCTTTCCAGTACTCTTGAGTCGGAACCCCGTCCGCTCCGAACG 1140  
Db |||||||  
1987 GACAACTCTTTCGGCGTCTTTCCAGTACTCTTGAGTCGGAACCCCGTCCGCTCCGAACG 2046  
QY 1141 AGATCCGTAATCTCGCGCGGAGGAGCACTGAGGAGTCGCAATCGACCGGATCGGAACACC 1200  
Db |||||||  
2047 AGATCCGTAATCTCGCGCGGAGGAGCACTGAGGAGTCGCAATCGACCGGATCGGAACACC 2106  
QY 1201 TCTCAGAAAGCGCTTAACAGTCACAGTCGCAAGATC 1239  
Db |||||||  
2107 TCTCAGAAAGCGCTTAACAGTCACAGTCGCAAGATC 2145

## RESULT 6

BD268253 8484 bp DNA linear PAT 17-JUL-2003  
LOCUS Adenovirus vector, packaging cell line, composition and method for  
DEFINITION production and use.  
ACCESSION BD268253  
VERSION BD268253.1 GI:33078021  
KEYWORDS JP 2002534130-A/57.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE other sequences; artificial sequences.  
AUTHORS 1 (bases 1 to 8484)  
Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and  
Sripichenko,Y.

TITLE Adenovirus vector, packaging cell line, composition and method for  
production and use  
JOURNAL Patent: JP 2002534130-A 57 15-OCT-2002;  
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE  
COMMENT OS Artificial Sequence  
PN JP 2002534130-A/57  
PD 15-OCT-2002

PF 14-JAN-2000 JP 2000593765  
PR 14-JAN-1999 US 60/115920  
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI  
SUSAN C STEVENSON,YELENA SKRIPCHENKO  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,  
PC C12N5/10,  
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC  
Description of Artificial Sequence: plasmid  
FH Key Location/Qualifiers  
FT source 1..8484  
FT /organism='Artificial Sequence'.  
FEATURES Location/Qualifiers  
source 1..8484  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

## ORIGIN

Query Match 99.9%; Score 1239; DB 6; Length 8484;  
Best Local Similarity 100.0%; Pred. No. 1.4e-237;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGATCCACTCTTCGGCATCGCTCTGCGAGGCGCAGCTGTTGGGTGAGTACTCCCT 60  
Db |||||||  
849 GGATCCACTCTTCGGCATCGCTCTGCGAGGCGCAGCTGTTGGGTGAGTACTCCCT 908  
QY 61 CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120  
Db |||||||  
909 CTGAAAAGCGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 968  
QY 121 GATATTACCTGGCGCGGCGTATGCGCTTTGAGGGTGGCGCATCCATCTGTCAGAAAA 180  
Db |||||||  
969 GATATTACCTGGCGCGGCGTATGCGCTTTGAGGGTGGCGCATCCATCTGTCAGAAAA 1028  
QY 181 GACAATCTTTTGTGTTCAAGCTTGGTGGCAACGACCGTAGAGGGCGTTGGACAGCAA 240

Db |||||||  
1029 GACAATCTTTTGTGTTCAAGCTTGGTGGCAACGACCCGTTAGAGGGCGTTGGACGAA 1088  
QY 241 CTTGGCGATGAGCGCAGGGGTTTGGTTTTTGTCCGATCGCGCGCTCTTTGGCGCGAT 300  
Db |||||||  
1089 CTTGGCGATGAGCGCAGGGGTTTGGTTTTTGTCCGATCGCGCGCTCTTTGGCGCGAT 1148  
QY 301 GTTTAGTGCACGTATTTCCGCGCAACGACCCGCTATTCGGGAAGACGGTGGTCCGCTC 360  
Db |||||||  
1149 GTTTAGTGCACGTATTTCCGCGCAACGACCCGCTATTCGGGAAGACGGTGGTCCGCTC 1208  
QY 361 GTCCGGCACCAAGGTGACGCGCAACCGCGTGTGTCAGAGGTGACAAGGTCAACGCTGGT 420  
Db |||||||  
1209 GTCCGGCACCAAGGTGACGCGCAACCGCGTGTGTCAGAGGTGACAAGGTCAACGCTGGT 1268  
QY 421 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGCA 480  
Db |||||||  
1269 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGCA 1328  
QY 481 GAATGGCGGTAGGGGTCTAGCTGCGTCTGTCGGGGGGTCTGCTCCAGCGTAAAGAC 540  
Db |||||||  
1329 GAATGGCGGTAGGGGTCTAGCTGCGTCTGTCGGGGGGTCTGCTCCAGCGTAAAGAC 1388  
QY 541 CCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGTCATCTTTGCAAGTCTAGCGCTG 600  
Db |||||||  
1389 CCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGTCATCTTTGCAAGTCTAGCGCTG 1448  
QY 601 CTGCCATCCGCGCGCGCAAGCGCGCTCGTATGGTGTGAGTGGGGAGCCCCATGGCAT 660  
Db |||||||  
1449 CTGCCATCCGCGCGCGCAAGCGCGCTCGTATGGTGTGAGTGGGGAGCCCCATGGCAT 1508  
QY 661 GGGTGGGTGAGCGCGAGGCGGTACATCGCGCAATGTCGTAAACGTAGAGGGGCTCTCT 720  
Db |||||||  
1509 GGGTGGGTGAGCGCGAGGCGGTACATCGCGCAATGTCGTAAACGTAGAGGGGCTCTCT 1568  
QY 721 GAGTATTTCCAAGATATGTAGGTTAGCATCTTCCACCGCGATGCTGGCGCGACGTAATC 780  
Db |||||||  
1569 GAGTATTTCCAAGATATGTAGGTTAGCATCTTCCACCGCGATGCTGGCGCGACGTAATC 1628  
QY 781 GTATAGTTCTGCGAGGAGCGAGGAGTCCGGACCGAGTTGTACGGGCGGCTGCTC 840  
Db |||||||  
1629 GTATAGTTCTGCGAGGAGCGAGGAGTCCGGACCGAGTTGTACGGGCGGCTGCTC 1688  
QY 841 TGCTCGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATATGTTGGACCGTG 900  
Db |||||||  
1689 TGCTCGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACCGTG 1748  
QY 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGAGGAGGCGGTAGGA 960  
Db |||||||  
1749 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGAGGAGGCGGTAGGA 1808  
QY 961 GTCCGCGAGCTGTTGACACGCTCGGGGTGACCTGCAAGTCTAGGGCGCAGTAGTCCAG 1020  
Db |||||||  
1809 GTCCGCGAGCTGTTGACACGCTCGGGGTGACCTGCAAGTCTAGGGCGCAGTAGTCCAG 1868  
QY 1021 GGTTCCTTGTATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080  
Db |||||||  
1869 GGTTCCTTGTATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1928  
QY 1081 GACAACTCTTTCGGCGTCTTTTCCAGTACTCTTTGATCGGAACCCGCTCGGCTCCGAACG 1140  
Db |||||||  
1929 GACAACTCTTTCGGCGTCTTTTCCAGTACTCTTTGATCGGAACCCGCTCGGCTCCGAACG 1988  
QY 1141 AGATCCGTAATCTCGCGCGGAGGCACTGAGGAGTCGCAATCGACCGGATCGGAACACC 1200  
Db |||||||  
1989 AGATCCGTAATCTCGCGCGGAGGCACTGAGGAGTCGCAATCGACCGGATCGGAACACC 2048  
QY 1201 TCTCAGAAAGCGCTTAACAGTCACAGTCGCAAGATC 1239  
Db |||||||  
2049 TCTCAGAAAGCGCTTAACAGTCACAGTCGCAAGATC 2087  
RESULT 7



A83180  
LOCUS A83180 10332 bp DNA linear PAT 21-JAN-2000  
DEFINITION Sequence 13 from Patent WO9851788.  
ACCESSION A83180  
VERSION A83180.1 GI:6732627  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 10332)  
AUTHORS Verheijen, J.H. and Quax, P.H.  
TITLE METHOD AND CONSTRUCT FOR INHIBITION OF CELL MIGRATION  
JOURNAL Patent: WO 9851788-A 13 19-NOV-1998;  
VERHEIJEN JOHAN HENDRIKUS (NL); TNO (NL)  
FEATURES  
Location/Qualifiers  
1..10332  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 91.5%; Score 1135; DB 6; Length 10332;  
Best Local Similarity 100.0%; Pred. No. 8.6e-217; Indels 0; Gaps 0;  
Matches 1135; Conservative 0; Mismatches 0;  
6 CACTCTCTCCGCATCGCTGCTGGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65  
4468 CACTCTCTCCGCATCGCTGCTGGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 4527  
66 AAGCGGGGATGATCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 125  
4528 AAGCGGGGATGATCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 4587  
126 TCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 185  
4588 TCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 4647  
186 TCTTTTGTGTCAAGCTTGTGGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 245  
4648 TCTTTTGTGTCAAGCTTGTGGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 4707  
246 CGATGAGCGCAGGCTTGTGTTTTGTGCGATGCGCGCGCTTCCTGCGCGGATGTTTA 305  
4708 CGATGAGCGCAGGCTTGTGTTTTGTGCGATGCGCGCGCTTCCTGCGCGGATGTTTA 4767  
306 GCTGCACTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 365  
4768 GCTGCACTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 4827  
366 GCACGAGGTGCACGCGCCAAACCGCGTGTGTCAGGCTGACAAAGGTCAACGCTGGTGCTA 425  
4828 GCACGAGGTGCACGCGCCAAACCGCGTGTGTCAGGCTGACAAAGGTCAACGCTGGTGCTA 4887  
426 CTTCTCCGCGTAGGCGCTGTTGGTCCAGACAGCGCGCGCTTCGCGCGAGCAATG 485  
4888 CTTCTCCGCGTAGGCGCTGTTGGTCCAGACAGCGCGCGCTTCGCGCGAGCAATG 4947  
486 GCGGTAGGGGTCTAGTCTGCTGTCGCGGGGGTCTGCTGTCACGGTAAAGACCCCGG 545  
4948 GCGGTAGGGGTCTAGTCTGCTGTCGCGGGGGTCTGCTGTCACGGTAAAGACCCCGG 5007  
546 GCAGCAGGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCAGTCTAGCGCTGCTGCC 605  
5008 GCAGCAGGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCAGTCTAGCGCTGCTGCC 5067  
606 ATGCGGGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGGT 665  
5068 ATGCGGGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGGT 5127  
666 GGGTGAGCGCGGAGCGGTACATGCGCGCAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 725  
5128 GGGTGAGCGCGGAGCGGTACATGCGCGCAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 5187

QY 726 TTCCAAGATATGTAGGTAGCATTTTCCACCGCGATGCTCGCGCACGTAATCGTATA 785  
Db 5188 TTCCAAGATATGTAGGTAGCATTTTCCACCGCGATGCTCGCGCACGTAATCGTATA 5247  
QY 786 GTTGTGCGAGGAGCGAGGAGTCCGACCGAGGTTGCTACGGCGGGCTGCTCTGCTC 845  
Db 5248 GTTGTGCGAGGAGCGAGGAGTCCGACCGAGTTCGACGGCGGGCTGCTCTGCTC 5307  
QY 846 GGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTGGACGCTGGAAGA 905  
Db 5308 GGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTGGACGCTGGAAGA 5367  
QY 906 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCAACGACGAGGAGCGTAGAGTCCG 965  
Db 5368 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCAACGACGAGGAGCGTAGAGTCCG 5427  
QY 966 GCAGCTTGTGACACGAGCTCGGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 1025  
Db 5428 GCAGCTTGTGACACGAGCTCGGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 5487  
QY 1026 CTTTGATGATGTCATATCTTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGGACAA 1085  
Db 5488 CTTTGATGATGTCATATCTTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGGACAA 5547  
QY 1086 ACTCTTCGCGTCTTTCCAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACG 1140  
Db 5548 ACTCTTCGCGTCTTTCCAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACG 5602  
RESULT 8  
BD082846  
LOCUS BD082846 10332 bp DNA linear PAT 27-AUG-2002  
DEFINITION Method and construct for inhibition of cell migration.  
ACCESSION BD082846  
VERSION BD082846.1 GI:22628456  
KEYWORDS JP 2001525669-A/13.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 10332)  
AUTHORS Quax, P.H.A. and Verheijen, J.H.  
TITLE Method and construct for inhibition of cell migration  
JOURNAL Patent: JP 2001525669-A 13 11-DEC-2001;  
NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK  
ONDERZOEK TNO  
COMMENT PN JP 2001525669-A/13  
PD 11-DEC-2001  
PF 11-MAY-1998 JP 1998549077  
PR 12-MAY-1997 EP 97201423.7  
PI PAULUS HUBERTUS ANDREAS QUAX, JOHAN HENDRIKUS VERHEIJEN PC  
C12N9/72, C12N15/62, C07K14/81//C07K19/00  
CC Strandedness: Unknown;  
CC Topology: Unknown;  
FH Key Location/Qualifiers.  
FEATURES  
source  
1..10332  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 91.5%; Score 1135; DB 6; Length 10332;  
Best Local Similarity 100.0%; Pred. No. 8.6e-217; Indels 0; Gaps 0;  
Matches 1135; Conservative 0; Mismatches 0;  
6 CACTCTCTCCGCATCGCTGCTGGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65  
4468 CACTCTCTCCGCATCGCTGCTGGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 4527  
66 AAGCGGGGATGATCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 125  
4528 AAGCGGGGATGATCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 4587  
126 TCACCTGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 185  
4588 TCACCTGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 4647  
186 TCTTTTGTGTCAAGCTTGTGGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 245  
4648 TCTTTTGTGTCAAGCTTGTGGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 4707  
246 CGATGAGCGCAGGCTTGTGTTTTGTGCGATGCGCGCGCTTCCTGCGCGGATGTTTA 305  
4708 CGATGAGCGCAGGCTTGTGTTTTGTGCGATGCGCGCGCTTCCTGCGCGGATGTTTA 4767  
306 GCTGCACTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 365  
4768 GCTGCACTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 4827  
366 GCACGAGGTGCACGCGCCAAACCGCGTGTGTCAGGCTGACAAAGGTCAACGCTGGTGCTA 425  
4828 GCACGAGGTGCACGCGCCAAACCGCGTGTGTCAGGCTGACAAAGGTCAACGCTGGTGCTA 4887  
426 CTTCTCCGCGTAGGCGCTGTTGGTCCAGACAGCGCGCGCTTCGCGCGAGCAATG 485  
4888 CTTCTCCGCGTAGGCGCTGTTGGTCCAGACAGCGCGCGCTTCGCGCGAGCAATG 4947  
486 GCGGTAGGGGTCTAGTCTGCTGTCGCGGGGGTCTGCTGTCACGGTAAAGACCCCGG 545  
4948 GCGGTAGGGGTCTAGTCTGCTGTCGCGGGGGTCTGCTGTCACGGTAAAGACCCCGG 5007  
546 GCAGCAGGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCAGTCTAGCGCTGCTGCC 605  
5008 GCAGCAGGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCAGTCTAGCGCTGCTGCC 5067  
606 ATGCGGGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGGT 665  
5068 ATGCGGGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGGT 5127  
666 GGGTGAGCGCGGAGCGGTACATGCGCGCAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 725  
5128 GGGTGAGCGCGGAGCGGTACATGCGCGCAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 5187



PUBMED  
REFERENCE  
10 (bases 10555 to 10732)  
Fowlkes,D.M. and Shenk,T.  
TITLE  
Transcriptional Control regions of the adenovirus VAI RNA gene  
JOURNAL  
Cell 22 (2 Pt 2), 405-413 (1980)  
MEDLINE  
81088343  
PUBMED  
7448868  
11 (bases 1653 to 4043)  
Bos,J.L., Polder,L.J., Bernards,R., Schrier,P.I., van den  
Elsen,P.J., van der Eb,A.J. and van Ormondt,H.  
TITLE  
The 2.2 kb Elb mRNA of human Ad12 and Ad5 codes for two tumor  
antigens starting at different AUG triplets  
JOURNAL  
Cell 27 (1 Pt 2), 121-131 (1981)  
MEDLINE  
82115327  
PUBMED  
7326748  
12 (bases 4001 to 6246)  
van Beveren,C.P., Maat,J., Dekker,B.M. and van Ormondt,H.  
TITLE  
The nucleotide sequence of the gene for protein IIVa2 and of the 5'  
leader segment of the major late mRNAs of adenovirus type 5  
JOURNAL  
Gene 16 (1-3), 179-189 (1981)  
MEDLINE  
82211779  
PUBMED  
7343420  
13 (bases 325 to 604)  
Hearing,P. and Shenk,T.  
TITLE  
Functional analysis of the nucleotide sequence surrounding the cap  
site for adenovirus type 5 region ElA messenger RNAs  
JOURNAL  
J. Mol. Biol. 167 (4), 809-822 (1983)  
MEDLINE  
83268691  
PUBMED  
6876165  
deletion mutants  
14 (bases 1 to 66)  
Nagata,K., Guggenheim,R.A. and Hurwitz,J.  
TITLE  
Specific binding of a cellular DNA replication protein to the  
origin of replication of adenovirus DNA  
Proc. Natl. Acad. Sci. U.S.A. 80 (20), 6177-6181 (1983)  
JOURNAL  
84016017  
PUBMED  
6336326  
15 (bases 6242 to 11570)  
Dekker,B.M. and van Ormondt,H.  
TITLE  
The nucleotide sequence of fragment HindIII-C of human adenovirus  
type 5 DNA (map positions 17.1-31.7)  
JOURNAL  
Gene 27 (1), 115-120 (1984)  
MEDLINE  
84183604  
PUBMED  
6325298  
16  
Downey,J.F., Eveleigh,C.M., Branton,P.E. and Bayley,S.T.  
TITLE  
Peptide maps and N-terminal sequences of polypeptides from early  
region 1A of human adenovirus 5  
J. Virol. 50 (1), 30-37 (1984)  
MEDLINE  
84138826  
PUBMED  
6699947  
sites; cds start for Elb proteins  
sites; splice sites in Elb 13S mRNA  
Notes on the presentation of ADENO in the EMBL data library: The  
genetic map of Adeno is customarily presented from left to right,  
the 0% position being left and the 100% position being right. The  
two strands of Adeno are normally represented like this: r-strand:  
3'-----5'  
l-strand: 5'-----3'  
0%  
This often causes confusion because the generally accepted way to  
represent DNA molecules is:  
5'-----3'  
3'-----5'  
Here Adeno virus sequences are always given in 5' to 3' direction  
and the sequence of the l-strand is displayed irrespective of the  
direction the viral transcription takes.  
This sequence corresponds to bases 1 to 11560 of <ad2>, which serve  
as some basis for the annotation of sites. The differences between  
<ad2> and <ad5> are too many to report herein, however a printout  
of those is available upon request from genbank. the map

coordinates in the sites presume 360 bases per map unit. although  
there are approximately 115 sequence differences between the two  
strains over this region, no site difference exceeds 0.02% by this  
calculation.  
the sequence represents the early mRNA transcripts elb and elb and  
the intermediate mRNA transcript ix, all of which are transcribed  
rightwardly off the r-strand; the IIVa2 and elb mRNAs which are  
transcribed leftwardly off the l-strand (indicated by '(c)' and  
'comp strand' below); and the 5' end of the 28 kb major late mRNA.  
the cap sites and possible promoter sequences for these are  
summarized in the following table:  
mRNA cap site possible promoter region -----  
-----  
tattata at bases 468-475 [6] elb 3582 1702  
tatataa at bases 1672-1678 [6] ix  
tatataa at bases 3551-3557 [6] IIVa2 5838 +/- 2 (c)  
tataaaa at bases 5979-5974 on the comp strand [10]  
major 6049 tataaaa at bases 6018-6024 [10] as  
with ad2, not all the transcripts from this region have been  
characterized at the sequence level. the nine proteins given in  
features table below are not the only possible gene products (see  
the main adenovirus 2 entry).  
large amounts of small RNAs are produced from the vai and vaii  
genes late in development for unknown reasons. [5] and [6]  
demonstrate that vai gene activity influences vai expression; that  
the 5' flank affects the start site of the rna but that an  
intragenic promoter (bases 10626 to 10690 below) determines whether  
the rna is actually produced; and that there is striking similarity  
between this rna and trna.  
FEATURES  
Location/Qualifiers  
1. .11570  
/organism="Human adenovirus type 5"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:28285"  
precursor\_RNA 499..1632  
/note="primary transcript of Elb region"  
mRNA join(499..1112,1229..1632)  
/note="mRNA 1 (part 1)"  
mRNA join(499..974,1229..1632)  
/note="mRNA 2 (part 1)"  
CDS join(560..1112,1229..1545)  
/note="unnamed protein product; Elb protein from 13S mRNA  
(32k, regulation and transformation)"  
/codon\_start=1  
/protein\_id="CAB40663.1"  
/db\_xref="GI:4584382"  
/db\_xref="GOA:P03255"  
/translation="MRHIICGGVITEEMASLLDQIEVLADNLPPSHPEPTLH  
ELYDLDTAPEDNEEAVSQIFPDVSMVLAQEGIDLLTFPPAGSPPPHLSRQEPQ  
EORALGPVSMPLNVEVIDLTCEAGFPSPDDEEGEEFVLDVVEPHGCRSHYH  
RRNVTGDPDIMCSLCYMTGCMFVYSPVSEPEPEPEPARPTRPKMAFALLRRPT  
Query Match 91.5%; Score 1135; DB 14; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 8.6e-217;  
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 CACTCTCTTCGGCATCGCTGTCTCGAGGGGCCAGCTGTTCGGGTGAGTACTCCCTCTCAA 65  
DB 6048 CACTCTCTTCGGCATCGCTGTCTCGAGGGGCCAGCTGTTCGGGTGAGTACTCCCTCTCAA 6107  
QY 66 AAGCGGGCATCTTCGGCTTAAGATTGTTCAGTTTCCAAAACAGGAGGATTGTAT 125  
DB 6108 AAGCGGGCATCTTCGGCTTAAGATTGTTCAGTTTCCAAAACAGGAGGATTGTAT 6167  
QY 126 TCACCTGCCCGCGGTGATGCTTTTGAGGGTGGCGGCATCCATCTGTGCAGAAAGACAA 185  
DB 6168 TCACCTGCCCGCGGTGATGCTTTTGAGGGTGGCGGCATCCATCTGTGCAGAAAGACAA 6227  
QY 186 TCTTTTGTCTGAAGCTTGTGGCAACGACCCGTAGAGGGCGTGTGACAGCACTTGG 245  
DB 6228 TCTTTTGTCTGAAGCTTGTGGCAACGACCCGTAGAGGGCGTGTGACAGCACTTGG 6287

QY 246 CGATGAGCGGAGGTTGGTTTTCGCGCATCGCGCGCTCCTTGGCGGATGTTTA 305  
DB 6288 CGATGAGCGGAGGTTGGTTTTCGCGCATCGCGCGCTCCTTGGCGGATGTTTA 6347  
QY 306 GCTGACAGTATTCGCGCGCAACGCGCATTCGGGAAAGACGGTGGTGGCTCGTGG 365  
DB 6348 GCTGACAGTATTCGCGCGCAACGCGCATTCGGGAAAGACGGTGGTGGCTCGTGG 6407  
QY 366 GCACAGGTGCACGCGCCCAACCGCGTGTGACAGGTGACAAAGTCAACGCTGGTGGCTA 425  
DB 6408 GCACAGGTGCACGCGCCCAACCGCGTGTGACAGGTGACAAAGTCAACGCTGGTGGCTA 6467  
QY 426 CCTCTCCGCTAGGCGCTCGTTGGTTCACGAGAGGCGCGCCCTTGGCGGACGAGATG 485  
DB 6468 CCTCTCCGCTAGGCGCTCGTTGGTTCACGAGAGGCGCGCCCTTGGCGGACGAGATG 6527  
QY 486 GCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCGTCCACGCTAAAGACCCCGG 545  
DB 6528 GCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCGTCCACGCTAAAGACCCCGG 6587  
QY 546 GCAGCAGCGCGCTCGAAGTAGTCTATCTTGCAATCTTGGCAATCTAGCGCTGCTGCC 605  
DB 6588 GCAGCAGCGCGCTCGAAGTAGTCTATCTTGCAATCTTGGCAATCTAGCGCTGCTGCC 6647  
QY 606 ATGCGCGCGCGCAACGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 665  
DB 6648 ATGCGCGCGCGCAACGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 6707  
QY 666 GGGTGAGCGCGGAGCGTACATCGCGCAAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 725  
DB 6708 GGGTGAGCGCGGAGCGTACATCGCGCAAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 6767  
QY 726 TTCCAAGATATAGGGTAGCATTTTCCACCGCGATCTGCGCGCACGCTAAATCGTATA 785  
DB 6768 TTCCAAGATATAGGGTAGCATTTTCCACCGCGATCTGCGCGCACGCTAAATCGTATA 6827  
QY 786 GTTCGTGCGAGGAGGAGGTGCGGACCGAGGTTCGTACGGCGGGGCTGCTGCTC 845  
DB 6828 GTTCGTGCGAGGAGGAGGTGCGGACCGAGGTTCGTACGGCGGGGCTGCTGCTC 6887  
QY 846 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTCGATGATGTTGGACGCTGGAAGA 905  
DB 6888 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTCGATGATGTTGGACGCTGGAAGA 6947  
QY 906 CTTTGAAGCTGCGCTGTGAGACTTACCGGTACGACGACGACGAGAGCGGTAGGAGTCGC 965  
DB 6948 CTTTGAAGCTGCGCTGTGAGACTTACCGGTACGACGACGACGAGAGCGGTAGGAGTCGC 7007  
QY 966 GCAGCTTGTGACCGAGCTCGCGGTGACCTGACAGCTTAGGGCGCAGTAGTCCAGGGTTT 1025  
DB 7008 GCAGCTTGTGACCGAGCTCGCGGTGACCTGACAGCTTAGGGCGCAGTAGTCCAGGGTTT 7067  
QY 1026 CTTTGAATGATGTCATATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1085  
DB 7068 CTTTGAATGATGTCATATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7127  
QY 1086 ACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGAAAG 1140  
DB 7128 ACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGAAAG 7182

RESULT 10  
Q0854904  
LOCUS Q0854904 31976 bp DNA linear PAT 23-AUG-2004  
DEFINITION Sequence 1 from Patent WO2004066947.  
ACCESSION Q0854904  
VERSION Q0854904.1 GI:51510464  
KEYWORDS . unidentified adenovirus  
SOURCE unidentified adenovirus  
ORGANISM Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
REFERENCE 1  
AUTHORS Hu, F. and Wu, B.

TITLE Therapy for primary and metastatic cancers  
JOURNAL Patent: WO 2004066947-A 1 12-AUG-2004;  
Shanghai Sunway Biotech Co Ltd (CN)  
FEATURES Location/Qualifiers  
source 1. 31976  
/organism="unidentified adenovirus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10535"  
ORIGIN  
Query Match 91.5%; Score 1135; DB 6; Length 31976;  
Best Local Similarity 100.0%; Pred. No. 7.9e-217;  
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 CACTCTCTTCGCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGGTACTCTCCCTCTGAA 65  
DB 5222 CACTCTCTTCGCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGGTACTCTCCCTCTGAA 5281  
QY 66 AAGCGGGCATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATTGATAT 125  
DB 5282 AAGCGGGCATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATTGATAT 5341  
QY 126 TCACCTCGCCCGCGCTCATGCTTGTAGGGTGGCGCATCCATCTGTGTCAGAAAGACAA 185  
DB 5342 TCACCTCGCCCGCGCTCATGCTTGTAGGGTGGCGCATCCATCTGTGTCAGAAAGACAA 5401  
QY 186 TCTTTTGTGTCAAGCTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 245  
DB 5402 TCTTTTGTGTCAAGCTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAACTTGG 5461  
QY 246 CGATGAGCGCGAGGGTTGGTTTGTGCGCATCGCGCGCTCCTTGGCGCGATGTTTA 305  
DB 5462 CGATGAGCGCGAGGGTTGGTTTGTGCGCATCGCGCGCTCCTTGGCGCGATGTTTA 5521  
QY 306 GCTCAGCTATTCCGCGCGCAACGACCGCATTCGGAAGACCGGTGGTGGCTCGTCGG 365  
DB 5522 GCTCAGCTATTCCGCGCGCAACGACCGCATTCGGAAGACCGGTGGTGGCTCGTCGG 5581  
QY 366 GCACAGGTGCACGCGCCAAACCGCGTTGTGACGGGTGACAAAGTCAACGCTGTGGCTA 425  
DB 5582 GCACAGGTGCACGCGCCAAACCGCGTTGTGACGGGTGACAAAGTCAACGCTGTGGCTA 5641  
QY 426 CTTCTCGCGTAGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTGGCGGACGAGATG 485  
DB 5642 CTTCTCGCGTAGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTGGCGGACGAGATG 5701  
QY 486 GCGGTAGGGGTCTAGCTGCGTCTGCTCCGGGGGTCTGCGTCCACGTAAGACCCCGG 545  
DB 5702 GCGGTAGGGGTCTAGCTGCGTCTGCTCCGGGGGTCTGCGTCCACGTAAGACCCCGG 5761  
QY 546 GCACGAGCGCGCTCGAAGTAGTCTATCTTGGCATCTTGGCAAGTCTAGCGCTGTGTC 605  
DB 5762 GCACGAGCGCGCTCGAAGTAGTCTATCTTGGCATCTTGGCAAGTCTAGCGCTGTGTC 5821  
QY 606 ATGCGCGCGCGCAAGCGCGCTCGTATGGTGGTGGGAGCCCGCATGGCATGGGGT 665  
DB 5822 ATGCGCGCGCGCAAGCGCGCTCGTATGGTGGTGGGAGCCCGCATGGCATGGGGT 5881  
QY 666 GGGTAGCGCGGAGCGCTACATGCGCAAAATGTCTGTAACGTCAGAGGGGTCTCTGAGTA 725  
DB 5882 GGGTAGCGCGGAGCGCTACATGCGCAAAATGTCTGTAACGTCAGAGGGGTCTCTGAGTA 5941  
QY 726 TTCCAAGATATGTTAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGCTAATCGTATA 785  
DB 5942 TTCCAAGATATGTTAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGCTAATCGTATA 6001  
QY 786 GTTCGTGCGAGGAGCGAGGAGGTTCGGACCGAGGTTCGTACGGCGGGCTGCTGCTC 845  
DB 6002 GTTCGTGCGAGGAGCGAGGAGGTTCGGACCGAGGTTCGTACGGCGGGCTGCTGCTC 6061  
QY 846 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTCGATGATGTTGGACGCTGGAAGA 905  
DB 6062 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTCGATGATGTTGGACGCTGGAAGA 6121

Qy	906	CGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTCAACGACAAAGAGGCGGTAGAGTCCG	965			
Db	6122	CGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTCAACGACAAAGAGGCGGTAGAGTCCG	6181			
Qy	966	GCAGCTTTGTTGACACAGCTCGGCGGTGACCTGCACGCTCTAGGGCGCAGTGTCCAGGGTTT	1025			
Db	6182	GCAGCTTTGTTGACACAGCTCGGCGGTGACCTGCACGCTCTAGGGCGCAGTGTCCAGGGTTT	6241			
Qy	1026	CTTTGATGATGTATATATATATATATATATATATATATATATATATATATATATATATATAT	1085			
Db	6242	CTTTGATGATGTATATATATATATATATATATATATATATATATATATATATATATATAT	6301			
Qy	1086	ACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACG	1140			
Db	6302	ACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAAACG	6356			
RESULT 11						
CQ854905						
LOCUS	CQ854905	31976 bp	DNA linear PAT 23-AUG-2004			
DEFINITION	Sequence 2 from Patent WO2004066947.					
ACCESSION	CQ854905					
VERSION	CQ854905.1	GI:51510465				
KEYWORDS						
SOURCE	unidentified adenovirus					
ORGANISM	unidentified adenovirus					
Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.						
REFERENCE	1					
AUTHORS	Hu, F. and Wu, B.					
TITLE	Therapy for primary and metastatic cancers					
JOURNAL	Patent: WO 2004066947-A 2 12-AUG-2004;					
Shanghai Sunway Biotech Co Ltd (CN)						
FEATURES	Location/Qualifiers					
source	1..31976					
/organism="unidentified adenovirus"						
/mol_type="unassigned DNA"						
/db_xref="taxon:10535"						
ORIGIN						
Query Match 91.5%; Score 1135; DB 6; Length 31976;						
Best Local Similarity 100.0%; Pred. No. 7.9e-217;						
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	6	CACCTCTCTTCGGCATCGCTGTCTCGAGGGCCAGCTGTGCGGTGAGTACTCCCTCTGAA	65			
Db	5222	CACCTCTCTTCGGCATCGCTGTCTCGAGGGCCAGCTGTGCGGTGAGTACTCCCTCTGAA	5281			
Qy	66	AAGCGGCGATGACTTCTCGCGCTAAGATGTGCTTTCCAAAACGAGAGGATTTGATAT	125			
Db	5282	AAGCGGCGATGACTTCTCGCGCTAAGATGTGCTTTCCAAAACGAGAGGATTTGATAT	5341			
Qy	126	TCACCTGGCCCGGCGGTGATGCTTTGAGGTGGCCGATCCATCTGTGTGAGAAAGACAA	185			
Db	5342	TCACCTGGCCCGGCGGTGATGCTTTGAGGTGGCCGATCCATCTGTGTGAGAAAGACAA	5401			
Qy	186	TCCTTTTGTGTTCAAGCTTGTGTGCAACGACCGTAGAGGGCTTGGACAGCAACTGG	245			
Db	5402	TCCTTTTGTGTTCAAGCTTGTGTGCAACGACCGTAGAGGGCTTGGACAGCAACTGG	5461			
Qy	246	CGATGAGCGCAGGGTTTGGTTTTTGTTCGGATCGGCGCTCCCTTGGCCGCGATGTTTA	305			
Db	5462	CGATGAGCGCAGGGTTTGGTTTTTGTTCGGATCGGCGCTCCCTTGGCCGCGATGTTTA	5521			
Qy	306	GCTGCAGTATTTCGCGGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTCGTCGG	365			
Db	5522	GCTGCAGTATTTCGCGGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTCGTCGG	5581			
Qy	366	GCACCAAGTGCACGCGCAACCGCGGTGTGACGGGTGACAAGTCAACCTCGTGGCTA	425			
Db	5582	GCACCAAGTGCACGCGCAACCGCGGTGTGACGGGTGACAAGTCAACCTCGTGGCTA	5641			
Qy	426	CCTCTCCGCTGAGCGCTCGTTGGTTCAGACAGAGGCGCGCCCTTTCGCGAGCAAGT	485			

```
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10,
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Adenovirus vector, packaging cell line,
composition and method
CC production and use
FH key Location/Qualifiers
FT source 1..32480
FT /organism='Adenovirus'.
Location/Qualifiers
1..32480
/organism='unidentified adenovirus'
/mol_type='genomic DNA'
/db_xref='taxon:10535'
FEATURES
source
ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 32480;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB 7192 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 7251
QY 66 AAGCGGCGATGACTTCTGCGGTAAAGATTGTCTGTTTCCAAAACAGAGGAGATTGATAT 125
DB 7252 AAGCGGCGATGACTTCTGCGGTAAAGATTGTCTGTTTCCAAAACAGAGGAGATTGATAT 7311
QY 126 TCACCTGCGCGCGGTGATGCTTTCGAGGTGGCGCATCCATCTGTCAGAAAGACAA 185
DB 7312 TCACCTGCGCGCGGTGATGCTTTCGAGGTGGCGCATCCATCTGTCAGAAAGACAA 7371
QY 186 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 245
DB 7372 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 7431
QY 246 CGATCGAGCGAGGCTGTTGTTTGTGCGCATCGCGCGTCTCTTGGCGCGGATGTTTA 305
DB 7432 CGATCGAGCGAGGCTGTTGTTTGTGCGCATCGCGCGTCTCTTGGCGCGGATGTTTA 7491
QY 306 GCTGCACGTAATTCGCGCGCAACGACCGCATTCGCGGAAGACGCTGTCGCTCGG 365
DB 7492 GCTGCACGTAATTCGCGCGCAACGACCGCATTCGCGGAAGACGCTGTCGCTCGG 7551
QY 366 GCACGAGTGCAACGCGCCAAACGCGGTGTGTCAGGCTGACAAAGTCAACGCTGTGGCTA 425
DB 7552 GCACGAGTGCAACGCGCCAAACGCGGTGTGTCAGGCTGACAAAGTCAACGCTGTGGCTA 7611
QY 426 CCTCTCCGCGTAGGCGCTGTTGTTGTCAGAGAGCGCGCGCTTGGCGGAGCAAGATG 485
DB 7612 CCTCTCCGCGTAGGCGCTGTTGTTGTCAGAGAGCGCGCGCTTGGCGGAGCAAGATG 7671
QY 486 GCGGTAGGGGCTAGCTGCTGCTCGCGGGGCTCTGCGTCCACGCTCAAGACCCCGG 545
DB 7672 GCGGTAGGGGCTAGCTGCTGCTCGCGGGGCTCTGCGTCCACGCTCAAGACCCCGG 7731
QY 546 GCAGCAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCGAAAGTCTAGCGCTGCTGCC 605
DB 7732 GCAGCAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCGAAAGTCTAGCGCTGCTGCC 7791
QY 606 ATGCGCGGCGCGCAACGCGCGCTCGTATGGGTTGAGTGGGGACCCCATATGGCATGGGGT 665
DB 7792 ATGCGCGGCGCGCAACGCGCGCTCGTATGGGTTGAGTGGGGACCCCATATGGCATGGGGT 7851
QY 666 GGGTCAGCGCGAGCGGTACATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 725
DB 7852 GGGTCAGCGCGAGCGGTACATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 7911
QY 726 TTCCAAAGATATAGGGTAGCATCTTTCACCGCGGATGTCGCGCGCACGCTAAATCGTATA 785
DB 7912 TTCCAAAGATATAGGGTAGCATCTTTCACCGCGGATGTCGCGCGCACGCTAAATCGTATA 7971
```

```
QY 786 GTTTCGTCGAGGAGCGAGGAGGCTCGGACCGAGGTTGCTTACGCGCGGCTGCTCTGCTC 845
DB 7972 GTTTCGTCGAGGAGCGAGGAGGCTCGGACCGAGGTTGCTTACGCGCGGCTGCTCTGCTC 8031
QY 846 GGAAGACTATCTGCTGCAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTTGGAGA 905
DB 8032 GGAAGACTATCTGCTGCAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTTGGAGA 8091
QY 906 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGACGAAAGAGCGGTAGAGTCGC 965
DB 8092 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGACGAAAGAGCGGTAGAGTCGC 8151
QY 966 GCAGCTTGTGACCGAGCTCGCGGTGACCTGACGCTTAGGGCCGAGTAGTCCAGGGTTT 1025
DB 8152 GCAGCTTGTGACCGAGCTCGCGGTGACCTGACGCTTAGGGCCGAGTAGTCCAGGGTTT 8211
QY 1026 CCTTGATGATGTATATCTTATCTGCTCCCTTTTTCACAGCTCGCGGTGAGGACAA 1085
DB 8212 CCTTGATGATGTATATCTTATCTGCTCCCTTTTTCACAGCTCGCGGTGAGGACAA 8271
QY 1086 ACTCTTCGCGGTCTTCCAGTACTCTTGGATCGAAACCCGTCGCGCTCCGAAACG 1140
DB 8272 ACTCTTCGCGGTCTTCCAGTACTCTTGGATCGAAACCCGTCGCGCTCCGAAACG 8326
RESULT 13
AR343138 32798 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 1 from patent US 6579522.
ACCESSION AR343138
VERSION AR343138.1 GI:33738640
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 32798)
AUTHORS Brough,D.E., King,C.R., Kovesdi,I. and Schaible,J.J.
TITLE Replication deficient adenoviral TNF vector
JOURNAL Patent: US 6579522-A 1 17-JUN-2003;
FEATURES
source
1..32798
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 32798;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB 5055 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 5114
QY 66 AAGCGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAACAGAGGAGATTGATAT 125
DB 5115 AAGCGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAACAGAGGAGATTGATAT 5174
QY 126 TCACCTGCGCGCGGTGATGCTTTCGAGGGTGGCGCATCCATCTGTCAGAAAGACAA 185
DB 5175 TCACCTGCGCGCGGTGATGCTTTCGAGGGTGGCGCATCCATCTGTCAGAAAGACAA 5234
QY 186 TCTTTTGTGTCAAGCTTGTGGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 245
DB 5235 TCTTTTGTGTCAAGCTTGTGGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 5294
QY 246 CGATGGAGCGCAGGTTGTTTGTTCGCGATCGCGCGCTCTCTGCGCGGATGTTTA 305
DB 5295 CGATGGAGCGCAGGTTGTTTGTTCGCGATCGCGCGCTCTCTGCGCGGATGTTTA 5354
QY 306 GCTGCACGTAATTCGCGCGCAACGACCGCATTCGCGGAAGACGCTGTCGCTCGG 365
DB 5355 GCTGCACGTAATTCGCGCGCAACGACCGCATTCGCGGAAGACGCTGTCGCTCGG 5414
```

QY 366 GCACACAGGTGACGCGCCCAACCGCGGTGTGTGACAGGGTGAACAAGTCAACGCTGGTGGCTA 425  
Db 5415 GCACACAGGTGACGCGCCCAACCGCGGTGTGTGACAGGGTGAACAAGTCAACGCTGGTGGCTA 5474  
QY 426 CTTCTCCGCGTAGGCGCTGTTGGTTCACAGAGCGCGCCCTTTGCGCGGAGCAGAAAG 485  
Db 5475 CTTCTCCGCGTAGGCGCTGTTGGTTCACAGAGCGCGCCCTTTGCGCGGAGCAGAAAG 5534  
QY 486 GCGTAGGAGGGGTCTAGCTGCGTCTCGTCCGCGGGGTCTGCGTCCACGGTAAAGACCCCGG 545  
Db 5535 GCGTAGGAGGGGTCTAGCTGCGTCTCGTCCGCGGGGTCTGCGTCCACGGTAAAGACCCCGG 5594  
QY 546 GCACAGGCGCGGTCCAAAGTAGTCTATCTTGATCTTGCATCTTGCAGTCTAGCGCTGCTGCC 605  
Db 5595 GCACAGGCGCGGTCCAAAGTAGTCTATCTTGATCTTGCATCTTGCAGTCTAGCGCTGCTGCC 5654  
QY 606 ATGCGCGGCGCAAGCGCGCTGATGGGTTCAGTGGGGGACCCCATGCGATGGGT 665  
Db 5655 ATGCGCGGCGCAAGCGCGCTGATGGGTTCAGTGGGGGACCCCATGCGATGGGT 5714  
QY 666 GGGTAGGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 725  
Db 5715 GGGTAGGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 5774  
QY 726 TTCCAAAGATATGATAGGTAAGCATCTTCCACCGCGATCTGCGCGCACGTAATCGTATA 785  
Db 5775 TTCCAAAGATATGATAGGTAAGCATCTTCCACCGCGATCTGCGCGCACGTAATCGTATA 5834  
QY 786 GTTCGTCGAGGCGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 845  
Db 5835 GTTCGTCGAGGCGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 5894  
QY 846 GGAAGACTATCTGCTGAAAGTGGATGCTGAGTGGATGATATGTTGGACGCTGGAAGA 905  
Db 5895 GGAAGACTATCTGCTGAAAGTGGATGCTGAGTGGATGATATGTTGGACGCTGGAAGA 5954  
QY 906 CGTTGAAGCTGCGGTCTGTGAGACTCTACCGGTTCACGACGAGGAGCGGTAGAGTCCG 965  
Db 5955 CGTTGAAGCTGCGGTCTGTGAGACTCTACCGGTTCACGACGAGGAGCGGTAGAGTCCG 6014  
QY 1026 CTTGATGATGTCATACTTATCTGTCCTTTTTCACAGCTCGCGGTTGAGACAA 1085  
Db 6075 CTTGATGATGTCATACTTATCTGTCCTTTTTCACAGCTCGCGGTTGAGACAA 6134  
QY 1086 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 1140  
Db 6135 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 6189

RESULT 14  
AX382187 32798 bp DNA linear PAT 18-MAR-2002  
LOCUS  
DEFINITION Sequence 1 from Patent WO0200906.  
ACCESSION AX382187  
VERSION AX382187.1 GI:19576990  
KEYWORDS  
SOURCE Human adenovirus type 5  
ORGANISM Human adenovirus type 5  
REFERENCE 1  
AUTHORS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
TITLES Brough, D.E., King, C.R. and Kovesdi, I.  
JOURNAL Replication deficient adenoviral tnf vector  
GENVEC, INC. (US) Patent: WO 020906-A 1 03-JAN-2002;

FEATURES  
Location/Qualifiers  
1..32798  
/organism="Human adenovirus type 5"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:28285"

ORIGIN  
Query Match 91.5%; Score 1135; DB 6; Length 32798;  
Best Local Similarity 100.0%; Pred. No. 7.9e-217;  
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 CACTCTCTTCGCGATCGTGTCTGAGGCGCAGCTGTTGGGTGAGTACTCCCTCTGAA 65  
Db 5055 CACTCTCTTCGCGATCGTGTCTGAGGCGCAGCTGTTGGGTGAGTACTCCCTCTGAA 5114  
QY 66 AAGCGGGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACAGAGGAGTATGATAT 125  
Db 5115 AAGCGGGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACAGAGGAGTATGATAT 5174  
QY 126 TCACCTGCGCCCGGTGATGCTTTGAGGGTGGCGCATCCCATCTGCTCAGAAAACACAA 185  
Db 5175 TCACCTGCGCCCGGTGATGCTTTGAGGGTGGCGCATCCCATCTGCTCAGAAAACACAA 5234  
QY 186 TCTTTTGTGTCAAGCTTGTGGCAACCGCCCTAGAGGCGTTGGACAGCAACTTGG 245  
Db 5235 TCTTTTGTGTCAAGCTTGTGGCAACCGCCCTAGAGGCGTTGGACAGCAACTTGG 5294  
QY 246 CGATGAGCGCAGGGTGTGTTTGTGCGCATCGCGCGCTCTTGGCGCGATGTTTA 305  
Db 5295 CGATGAGCGCAGGGTGTGTTTGTGCGCATCGCGCGCTCTTGGCGCGATGTTTA 5354  
QY 306 GCTGCACGTATTCGCGCGCAACGCGCATTCGGGAAAGACGTTGGTGGCTCGTCGG 365  
Db 5355 GCTGCACGTATTCGCGCGCAACGCGCATTCGGGAAAGACGTTGGTGGCTCGTCGG 5414  
QY 366 GCACAGGTGCAACCGCCCAACCGCGTGTGTCAGGGTGAACAAGTCAACGCTGGTGGCTA 425  
Db 5415 GCACAGGTGCAACCGCCCAACCGCGTGTGTCAGGGTGAACAAGTCAACGCTGGTGGCTA 5474  
QY 426 CCTCTCCGCTAGGCGCTGTTGGTTCAGAGAGCGCGCGCTTGGCGGAGCAAGT 485  
Db 5475 CCTCTCCGCTAGGCGCTGTTGGTTCAGAGAGCGCGCGCTTGGCGGAGCAAGT 5534  
QY 486 GGGTAGGGGTCTAGCTGCGTCTCGTCCGCGGGTCTGCGTCCACGGTAAAGACCCCGG 545  
Db 5535 GGGTAGGGGTCTAGCTGCGTCTCGTCCGCGGGTCTGCGTCCACGGTAAAGACCCCGG 5594  
QY 546 GCACAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTGGCAAGTCTAGCGCTGCTGCC 605  
Db 5595 GCACAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTGGCAAGTCTAGCGCTGCTGCC 5654  
QY 606 ATGCGCGGCGCAAGCGCGCTGATGGGTGAGTGGGGGACCCCATGCGATGGGT 665  
Db 5655 ATGCGCGGCGCAAGCGCGCTGATGGGTGAGTGGGGGACCCCATGCGATGGGT 5714  
QY 666 GGGTAGGCGGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 725  
Db 5715 GGGTAGGCGGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 5774  
QY 726 TTCCAAAGATATGATAGGTAAGCATCTTCCACCGCGATCTGCGCGCACGTAATCGTATA 785  
Db 5775 TTCCAAAGATATGATAGGTAAGCATCTTCCACCGCGATCTGCGCGCACGTAATCGTATA 5834  
QY 786 GTTCGTCGAGGCGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 845  
Db 5835 GTTCGTCGAGGCGAGGAGTCCGGACCGAGGTGCTACGGGGGCTCTCTGCTC 5894  
QY 846 GGAAGACTATCTGCTGAAAGTGGATGCTGAGTGGATGATATGTTGGACGCTGGAAGA 905  
Db 5895 GGAAGACTATCTGCTGAAAGTGGATGCTGAGTGGATGATATGTTGGACGCTGGAAGA 5954  
QY 906 CGTTGAAGCTGCGGTCTGTGAGACTCTACCGGTTCACGACGAGGAGCGGTAGAGTCCG 965  
Db 5955 CGTTGAAGCTGCGGTCTGTGAGACTCTACCGGTTCACGACGAGGAGCGGTAGAGTCCG 6014  
QY 966 GCAGCTTGTGACACGCTCGCGGTGACCTGCAAGTCTAGGCGCGAGTAGTCCAGGGTTT 1025  
Db 6015 GCAGCTTGTGACACGCTCGCGGTGACCTGCAAGTCTAGGCGCGAGTAGTCCAGGGTTT 6074



```
QY 1026 CCTTGATGATGTCATACCTATCTCTCCCTTTTTCACAGCTCGCGTTGAGGCAA 1085
Db 6075 CCTTGATGATGTCATACCTATCTCTCCCTTTTTCACAGCTCGCGTTGAGGCAA 6134
QY 1086 ACTCTTCGCGGTCTTTCAGTACTCTTGATCGGAACCGCTCGCGCTCCGAAACG 1140
Db 6135 ACTCTTCGCGGTCTTTCAGTACTCTTGATCGGAACCGCTCGCGCTCCGAAACG 6189

RESULT 15
CQ854906 32802 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 3 from Patent WO2004066947.
ACCESSION CQ854906
VERSION CQ854906.1 GI:51510466
KEYWORDS unidentified adenovirus
SOURCE unidentified adenovirus
ORGANISM unidentified adenovirus
REFERENCE 1
AUTHORS Hu.F. and Wu.B.
TITLE Therapy for primary and metastatic cancers
JOURNAL Patent: WO 2004066947-A 3 12-AUG-2004;
Shanghai Sunway Biotech Co Ltd (CN)
FEATURES
    source
        1..32802
            /organism="unidentified adenovirus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:10535"

ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 32802;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CACTCTTCGCGATCGCTGTCTGCGAGGCCAGCTGTGGGTGAGTACTCCCTCTGAA 65
Db 6048 CACTCTTCGCGATCGCTGTCTGCGAGGCCAGCTGTGGGTGAGTACTCCCTCTGAA 6107

QY 66 AAGCGGGCATGATCTTCGCGTAAAGTTGTTCAGTTTCCAAAACAGGAGGATTTGATAT 125
Db 6108 AAGCGGGCATGATCTTCGCGTAAAGTTGTTCAGTTTCCAAAACAGGAGGATTTGATAT 6167

QY 126 TCACCTGGCCCGCGGTGATGCTTTGAGGTGGCCGATCCATCTGGTCAAGAAAGACAA 185
Db 6168 TCACCTGGCCCGCGGTGATGCTTTGAGGTGGCCGATCCATCTGGTCAAGAAAGACAA 6227

QY 186 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGTTGGACAGCAACTTGG 245
Db 6228 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGTTGGACAGCAACTTGG 6287

QY 246 CGATGGAGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTCTTGGCCGCGATGTTTA 305
Db 6288 CGATGGAGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTCTTGGCCGCGATGTTTA 6347

QY 306 GGTGCAAGTATTCGCGCGCAACGACCCGATTCGGGAAAGACGTTGGTGGCTCGTCGG 365
Db 6348 GGTGCAAGTATTCGCGCGCAACGACCCGATTCGGGAAAGACGTTGGTGGCTCGTCGG 6407

QY 366 GCACAGGTGACGCGCCAAACCGCGTGTGACAGGTGACAGGTCAACGCTGGTGCTA 425
Db 6408 GCACAGGTGACGCGCCAAACCGCGTGTGACAGGTGACAGGTCAACGCTGGTGCTA 6467

QY 426 CCTCTCCGCTAGGCGCTGTTGGTCCAGCAGAGCGCGCGCTTGGCGGAGCAAGATG 485
Db 6468 CCTCTCCGCTAGGCGCTGTTGGTCCAGCAGAGCGCGCGCTTGGCGGAGCAAGATG 6527

QY 486 GCGGTAGGGGGTCTAGTCTGCTCTGTCGGGGGGTCTGCTGTCACGGTAAGACCCCGG 545
Db 6528 GCGGTAGGGGGTCTAGTCTGCTCTGTCGGGGGGTCTGCTGTCACGGTAAGACCCCGG 6587

QY 546 GCAGCAGGCGCGTCGAAGTAGTCTATCTTCATCTCTTGCAAGTCTAGCGCTCTGCTGCC 605
```

```
Db 6588 GCACAGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCATCTTGCATCTAGCGCTCTGCC 6647
QY 606 ATGCGCGCGCGCAAGCGCGCTCGTATGAGTTGAGTGGGGGACCCCATGGCATGGGGT 665
Db 6648 ATGCGCGCGCGCAAGCGCGCTCGTATGAGTTGAGTGGGGGACCCCATGGCATGGGGT 6707
QY 666 GGGTGAAGCGGAGCGCTACATGCGGCAAAATGCTAAACCTAGAGGGGCTCTCTGAGTA 725
Db 6708 GGGTGAAGCGGAGCGCTACATGCGGCAAAATGCTAAACCTAGAGGGGCTCTCTGAGTA 6767
QY 726 TTCCAAGATATGATAGGTAGCATCTTCCACCGCGGATCTCGCGCGCACGTAATCGTATA 785
Db 6768 TTCCAAGATATGATAGGTAGCATCTTCCACCGCGGATCTCGCGCGCACGTAATCGTATA 6827
QY 786 GTTCGTGCGAGGAGCGAGGAGGTGCGGACCCGAGTTGCTACGCGCGGGCTCTCTGCTC 845
Db 6828 GTTCGTGCGAGGAGCGAGGAGGTGCGGACCCGAGTTGCTACGCGCGGGCTCTCTGCTC 6887
QY 846 GGAAGACTATCTGCTGAAGATGTCATGTCGATGAGTTGGATGATATGTTGGACGCTGGAAGA 905
Db 6888 GGAAGACTATCTGCTGAAGATGTCATGTCGATGAGTTGGATGATATGTTGGACGCTGGAAGA 6947
QY 906 CGTTGAAGCTGCGCTCTGTGAGACCTACCGCGTCAACGCAAGAGGCGTAGGAGTCCG 965
Db 6948 CGTTGAAGCTGCGCTCTGTGAGACCTACCGCGTCAACGCAAGAGGCGTAGGAGTCCG 7007
QY 966 GCAGCTTGTGACCCAGCTCGCGCGGTGACCTGCAAGTCTGAGGCGGAGTAGTCCAGGGTTT 1025
Db 7008 GCAGCTTGTGACCCAGCTCGCGCGGTGACCTGCAAGTCTGAGGCGGAGTAGTCCAGGGTTT 7067
QY 1026 CCTTGATGATGTCATACCTATCTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGCGCTCCGAAACG 1140
Db 7068 CCTTGATGATGTCATACCTATCTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGCGCTCCGAAACG 7127
QY 1086 ACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGCGCTCCGAAACG 1140
Db 7128 ACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGCGCTCCGAAACG 7182
```

Search completed: July 14, 2005, 14:03:29  
Job time : 9291.17 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 1748.26 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-32  
Perfect score: 1240  
Sequence: 1 ggcattcactctctccgcac.....cagtcacagtcgcaagatct 1240

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0  
Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1240	100.0	1240	3	AAA59060 Nucleotid
2	1240	100.0	1240	6	ABA94272 Adenoviru
3	1240	100.0	1240	10	ADB75118 Adenoviru
4	1240	100.0	1240	10	ADF48742 Ad5 tripa
5	1239	99.9	7231	3	AAA59090 Nucleotid
6	1239	99.9	7231	3	ABA94286 Nucleotid
7	1239	99.9	7231	10	ADF75132 Plasmid p
8	1239	99.9	7231	10	ADF48774 Adenoviru
9	1239	99.9	7960	3	AAA59072 Nucleotid
10	1239	99.9	7960	6	ABA94274 Nucleotid
11	1239	99.9	7960	10	ADB75120 Plasmid p
12	1239	99.9	7960	10	ADF48754 Fibre exp
13	1239	99.9	7989	3	AAA59075 Nucleotid
14	1239	99.9	7989	6	ABA94277 Nucleotid
15	1239	99.9	7989	10	ADB75123 Plasmid p
16	1239	99.9	7989	10	ADF48757 Fibre exp
17	1239	99.9	8383	3	AAA59071 Nucleotid
18	1239	99.9	8383	6	ABA94273 Nucleotid
19	1239	99.9	8383	10	ADB75119 Plasmid p
20	1239	99.9	8383	10	ADF48753 Fibre exp

21	1239	99.9	8484	3	AAA59091 Nucleotid
22	1239	99.9	8484	10	ADF48775 Fibre exp
23	1135	91.5	10332	2	AAV33921 Nucleotid
24	1135	91.5	31183	4	AA003963 Adenoviru
25	1135	91.5	31446	3	AAA09088 AdPB-beta
26	1135	91.5	31880	12	ADO09305 Ad09305 WTI-F-ade
27	1135	91.5	31976	13	ADRA1670 Oncolytic
28	1135	91.5	31976	13	ADRA1669 Adr41669 Oncolytic
29	1135	91.5	32026	2	AAT60559 Recombina
30	1135	91.5	32165	3	AAA09092 AdMTV-be
31	1135	91.5	32165	3	AAA14723 Nucleotid
32	1135	91.5	32166	3	AAA09090 AdPSA-bet
33	1135	91.5	32166	4	AAAC89170 AdRSVPHYD
34	1135	91.5	32167	3	AAA14803 Nucleotid
35	1135	91.5	32167	3	AAZ93332 Partial s
36	1135	91.5	32409	12	ADO36637 Adenovira
37	1135	91.5	32480	3	AAA59055 Nucleotid
38	1135	91.5	32480	6	ABA94267 Adenoviru
39	1135	91.5	32480	10	ADB75113 Adenovira
40	1135	91.5	32480	10	ADF48737 Adenovira
41	1135	91.5	32681	12	ADO36636 Adenovira
42	1135	91.5	32798	6	ABA97684 Replicati
43	1135	91.5	32802	13	ADR41671 Adr41671 S98-100 (
44	1135	91.5	32886	3	AAA09086 AdRSV-bet
45	1135	91.5	33014	13	ADP79484 Adenoviru

ALIGNMENTS

RESULT 1  
AAA59060  
ID AAA59060 standard; DNA; 1240 BP.  
XX  
AC AAA59060;

XX 15-SEP-2003 (revised)  
DT 07-NOV-2000 (first entry)

XX Nucleotide sequence of a tripartite leader sequence.

XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;  
KW ss.

XX Human adenovirus type 5.

XX WO200042208-A1.

XX 20-JUL-2000.

XX 14-JAN-2000; 2000WO-EP000265.

XX 14-JAN-1999; 99US-0115920P.

XX (NOVS ) NOVARTIS AG.

XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX (SCHI ) SCRIPPS RES INST.

XX Nemerow GR, Von Seggern DU, Hallenbeck PL, Stevenson SC;

XX Skripchenko Y;

XX WPI; 2000-476068/41.

XX New nucleic acid comprising an adenovirus tripartite leader nucleotide  
PT for producing high-capacity and targeted vectors for adenovirus-based  
PT gene therapy.

XX Claim 5; Page 180; 212pp; English.

XX The specification describes a nucleic acid molecule comprising an  
XX adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence  
XX comprising two different TPL exons or three same or different TPL exons.  
XX The nucleic acid is used to produce an adenovirus vector particle,

CC deliver an exogenous gene to a target cell, pseudotype recombinant viral  
CC vectors, target an adenovirus vector to a cell, produce a modified  
CC adenovirus, deliver a heterologous gene to an animal and produce a  
CC gutless adenoviral vector particle. The present sequence represents a TPL  
CC sequence, which is used to construct nucleic acid molecules of the  
CC invention. (Updated on 15-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1240; DB 3; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGATCCACTCTCTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 60  
DB 1 GGATCCACTCTCTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 60  
QY 61 CTGAAGAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATT 120  
DB 61 CTGAAGAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATT 120  
QY 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAAAAA 180  
DB 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAAAAA 180  
QY 181 GACAATCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGGTTGGACAGCA 240  
DB 181 GACAATCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGGTTGGACAGCA 240  
QY 241 CTTGGCGATGAGCGCAGGGTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGAT 300  
DB 241 CTTGGCGATGAGCGCAGGGTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGAT 300  
QY 301 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTCCGGAAGACGGTGGTGGCTC 360  
DB 301 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTCCGGAAGACGGTGGTGGCTC 360  
QY 361 GTGCGGACAGGTGACGCGCAACCGGGTTGTGAGGTGACAGGTCAACGCTGGT 420  
DB 361 GTGCGGACAGGTGACGCGCAACCGGGTTGTGAGGTGACAGGTCAACGCTGGT 420  
QY 421 GGCTACCTCTCCGCTAGCGCTCGTTGTTCCAGAGAGGGCGCCCTTGGCGGACA 480  
DB 421 GGCTACCTCTCCGCTAGCGCTCGTTGTTCCAGAGAGGGCGCCCTTGGCGGACA 480  
QY 481 GAATGGCGGTAGGGGTTCTAGTGGCTCTCGTCCGGGGGGTCTGCTCCACGGTAAAGAC 540  
DB 481 GAATGGCGGTAGGGGTTCTAGTGGCTCTCGTCCGGGGGGTCTGCTCCACGGTAAAGAC 540  
QY 541 CCGGGGACGAGCGCGCGCTGAAAGTAGTCTATCTTGCATCTTGCAGAGCTAGGCGCTG 600  
DB 541 CCGGGGACGAGCGCGCGCTGAAAGTAGTCTATCTTGCATCTTGCAGAGCTAGGCGCTG 600  
QY 601 CTGCCATGCGGGCGGCAAGCGCGCTCGTATGAGTTGAGTGGGGACCCCATGGAT 660  
DB 601 CTGCCATGCGGGCGGCAAGCGCGCTCGTATGAGTTGAGTGGGGACCCCATGGAT 660  
QY 661 GGGGTGGGTAGCGGGAGCGGTACATGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 720  
DB 661 GGGGTGGGTAGCGGGAGCGGTACATGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 720  
QY 721 GAGTATTCAGATATGATAGGTAGCATCTTCCACGGGATGCTGGCGCGCATATC 780  
DB 721 GAGTATTCAGATATGATAGGTAGCATCTTCCACGGGATGCTGGCGCGCATATC 780  
QY 781 GTATAGTTCTGCGAGGGAGCGAGGTCGGAGCCGAGGTTGCTACGGGGGCTGCTC 840  
DB 781 GTATAGTTCTGCGAGGGAGCGAGGTCGGAGCCGAGGTTGCTACGGGGGCTGCTC 840  
QY 841 TGCTCGGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900  
DB 841 TGCTCGGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900

QY 901 GAACACCTTGAAGCTGGGCTGTGTGAGACCTACCGCTCAGCAGAGGAGCGGTAGGA 960  
DB 901 GAACACCTTGAAGCTGGGCTGTGTGAGACCTACCGCTCAGCAGAGGAGCGGTAGGA 960  
QY 961 GTCCGCGACAGCTTGTGTGACCAAGCTCGGGGTGACCTGACAGCTTAGGGCGCAGTAGTCCAG 1020  
DB 961 GTCCGCGACAGCTTGTGTGACCAAGCTCGGGGTGACCTGACAGCTTAGGGCGCAGTAGTCCAG 1020  
QY 1021 GGTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACAGCTCGCGTTGAG 1080  
DB 1021 GGTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACAGCTCGCGTTGAG 1080  
QY 1081 GACAAACTCTTCGCGCTTTTCCAGTACTCTTGTGATCGGAACCGCTCGCGCTCCGAACG 1140  
DB 1081 GACAAACTCTTCGCGCTTTTCCAGTACTCTTGTGATCGGAACCGCTCGCGCTCCGAACG 1140  
QY 1141 AGATCCGCTACTCCCGCCCGGAGGACCTGACGAGTCCGATCGACCGGATCGGAACG 1200  
DB 1141 AGATCCGCTACTCCCGCCCGGAGGACCTGACGAGTCCGATCGACCGGATCGGAACG 1200  
QY 1201 TCTCGAAGAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240  
DB 1201 TCTCGAAGAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240  
RESULT 2  
ID ABA94272 standard; DNA; 1240 BP.  
XX ABA94272;  
XX  
XX 07-AUG-2003 (revised)  
DT 13-MAR-2002 (first entry)  
XX  
DE Adenovirus 5 tripartite leader (TPL) nucleotide sequence.  
XX  
KW Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; scargard disease gene; STDG1;  
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; tripartite leader; TPL; da.  
XX Human adenovirus type 5.  
XX  
XX WO200183729-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 30-APR-2001; 2001WO-EP004863.  
XX  
XX 01-MAY-2000; 2000US-00562934.  
XX  
XX (NOVS) NOVARTIS AG.  
XX (SCRI) SCRIPPS RES INST.  
XX (NEME) NEMEROW G R.  
XX (VSEG) VON SEGGERN D J.  
XX (FRIE) FRIEDLANDER M.  
XX  
XX Nemerow GR, Von Seggern DJ, Friedlander M;  
XX WPI; 2002-082846/11.  
XX  
XX Polynucleotide for making vectors, useful for treating ocular diseases,  
XX e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat  
XX sequences, packaging signal and photoreceptor-specific promoter.  
XX  
XX Example 3; Page 131-132; 149pp; English.  
XX  
XX The invention provides an isolated polynucleotide comprising adenovirus  
XX (AV) inverter terminal repeat sequences (ITRS), AV packaging signal  
XX operatively linked to ITRS and a photoreceptor-specific promoter. A  
XX recombinant AV vector (AVV) comprising the polynucleotide is useful for  
XX targeted delivery of a gene product to the eye (especially to the  
XX vitreous cavity), for treating an ocular disease, e.g., retinal

CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic  
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal  
CC preferably human. The AAV comprises a fiber protein that specifically or  
CC selectively binds to receptors that are expressed on cells (preferably  
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a  
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein  
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5  
CC penton, and the therapeutic product is a tropic factor, an anti-  
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type  
CC stargardt disease gene (STGD1), an anti-cancer agent and a protein that  
CC regulates expression of a photoreceptor specific gene product. The viral  
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV  
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful  
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber  
CC protein or its portion, and selectively transduces photoreceptors and  
CC delivers a gene product encoded by AAV. The present sequence represents a  
CC adenovirus 5 tripartite leader (TPL) nucleotide sequence. (Updated on 07-  
XX AUG-2003 to correct OS field.)

Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 1240; DB 6; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATCGCTCTCTCGAGGSCCAGCTGTGGGGTGAATCTCCCT 60  
DB 1 GGATCCACTCTCTCCGATCGCTCTCTCGAGGSCCAGCTGTGGGGTGAATCTCCCT 60  
QY 61 CTGAAAGGGGCGATGCTTCCGCTAAGATTGTCTGTTCCAAAACGAGGAGATT 120  
DB 61 CTGAAAGGGGCGATGCTTCCGCTAAGATTGTCTGTTCCAAAACGAGGAGATT 120  
QY 121 GATATTCACCTGGCCGCGGTGATCCCTTTGAGGGTGGCCGATCATCTGTGTGAGAAA 180  
DB 121 GATATTCACCTGGCCGCGGTGATCCCTTTGAGGGTGGCCGATCATCTGTGTGAGAAA 180  
QY 181 GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGGTGGACAGCAA 240  
DB 181 GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGGTGGACAGCAA 240  
QY 241 CTGGCGATGAGCGAGGGTTGGTTTTGTGCGATCGGCGCGCTCTTGGCGCGCAT 300  
DB 241 CTGGCGATGAGCGAGGGTTGGTTTTGTGCGATCGGCGCGCTCTTGGCGCGCAT 300  
QY 301 GTTTAGCTGCAGTATTCGCGCAACGACCGCCATTCGGAAGACGGTGTGCGCTC 360  
DB 301 GTTTAGCTGCAGTATTCGCGCAACGACCGCCATTCGGAAGACGGTGTGCGCTC 360  
QY 361 GTGCGGCACCAAGTGCAGCGCCAAACCGCGTGTGTCAGGGTGAACAAGTCAACCGTGT 420  
DB 361 GTGCGGCACCAAGTGCAGCGCCAAACCGCGTGTGTCAGGGTGAACAAGTCAACCGTGT 420  
QY 421 GGTACTCTCTCCGTAGCGCTCGTGTGTCAGAGAGCGCGCCCTTGGCGGAGCA 480  
DB 421 GGTACTCTCTCCGTAGCGCTCGTGTGTCAGAGAGCGCGCCCTTGGCGGAGCA 480  
QY 481 GAAATGCGGTAGGGGTCTAGTGTGCTCTGTCGCGGGGGTCTGCTCCACGGTAAAGAC 540  
DB 481 GAAATGCGGTAGGGGTCTAGTGTGCTCTGTCGCGGGGGTCTGCTCCACGGTAAAGAC 540  
QY 541 CCCGGGACAGCGCGCTGCAAGTGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 600  
DB 541 CCCGGGACAGCGCGCTGCAAGTGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 600  
QY 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660  
DB 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660  
QY 661 GGGGTGGGTAGCGCGGAGGCGTATGCGCGAAATGTCGTAACGTAAGAGGGGTCTCT 720  
DB 661 GGGGTGGGTAGCGCGGAGGCGTATGCGCGAAATGTCGTAACGTAAGAGGGGTCTCT 720

QY 721 GAGTATTTCCAAAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 780  
DB 721 GAGTATTTCCAAAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 780  
QY 781 GTATAGTTCCGCGAGGAGGAGAGGTCCGGACCGAGGTTGCTACCGGGGGGTGCTC 840  
DB 781 GTATAGTTCCGCGAGGAGGAGAGGTCCGGACCGAGGTTGCTACCGGGGGGTGCTC 840  
QY 841 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900  
DB 841 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900  
QY 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTACGACGAGGAGGCGTAGGA 960  
DB 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTACGACGAGGAGGCGTAGGA 960  
QY 961 GTGCGGAGGCTGTTGTTGACGAGCTGGGCGGTGACCTGTAGGCGGCGAGTAGTCCAG 1020  
DB 961 GTGCGGAGGCTGTTGTTGACGAGCTGGGCGGTGACCTGTAGGCGGCGAGTAGTCCAG 1020  
QY 1021 GGTTCCTTGTATGATGTATCATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080  
DB 1021 GGTTCCTTGTATGATGTATCATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080  
QY 1081 GACAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAAACG 1140  
DB 1081 GACAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAAACG 1140  
QY 1141 AGATCCGTAACCTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 1200  
DB 1141 AGATCCGTAACCTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 1200  
QY 1201 TCTCGAAGAAAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240  
DB 1201 TCTCGAAGAAAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240

## RESULT 3

ADB75118

ID ADB75118 standard; DNA; 1240 BP.

XX ADB75118;

XX 04-DEC-2003 (first entry)

XX Adenovirus type 5 tripartite leader sequence #2.

XX ophthalmological; antiinflammatory; antidiabetic; gene therapy;

KW adenovirus inverted terminal repeat sequence;

KW adenovirus packaging signal; photoreceptor-specific promoter;

KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;

KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;

KW rhodopsin; wild-type Stargardt disease gene; STGD1; anti-cancer agent;

KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;

KW gyrate atrophy; macular dystrophy; retinoblastoma;

KW photoreceptor-restricted transgene expression;

KW recombinant adenovirus vector; adenovirus type 5; Ad5;

KW tripartite leader sequence; TPL; ds.

XX Human adenovirus type 5.

XX US2002193327-A1.

XX 19-DEC-2002.

XX 01-MAY-2001; 2001US-00847101.

XX 01-MAY-2000; 2000US-00562934.

XX (SCRI ) SCRIPPS RES INST.

XX Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2003-657234/62.

XX Novel nucleic acids comprising adenovirus inverted terminal repeat

PT sequences, adenovirus packaging signals operatively linked to the

PT sequences and photoreceptor-specific promoters, useful for treating

PT retinitis pigmentosa.

XX

PS Example 3; Page 78; 106pp; English.

XX

CC The invention describes an isolated nucleic acid (I) comprising

CC adenovirus inverted terminal repeat sequence, an adenovirus packaging

CC signal operatively linked to the sequence, and a photoreceptor-specific

CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful

CC for targeted delivery of a gene product to the eye of a mammal which

CC involves administering (II) that comprises heterologous DNA encoding the

CC gene product or resulting in expression of the gene product, where the

CC recombinant virus comprises a fibre protein that specifically or

CC selectively binds to receptors that are expressed on cells which are

CC photoreceptors, in the eye. The recombinant virus comprises a fibre

CC protein which is an adenovirus type 37, from an adenovirus type D

CC serotype. The fibre is a chimeric protein containing a sufficient portion

CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for

CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient

CC portion of an adenovirus serotype D knob portion of the fiber for

CC selective binding to photoreceptors in the eye of a mammal. The

CC encapsulated nucleic acid comprises a photoreceptor-specific promoter

CC operatively linked to a nucleic acid comprising the therapeutic product

CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding

CC a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-

CC cancer agent and a protein that regulates expression of a photoreceptor-

CC specific gene product. The delivery is effected for treatment of an

CC ocular disease such as retinal degenerative disease e.g., retinitis

CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal

CC vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or

CC retinoblastoma inherited and acquired retinal and neovascular

CC degenerative diseases. The viral nucleic acid comprises an adenovirus

CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging

CC signal operatively linked to the sequence. The ITRs and packaging signal

CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or

CC 5. The viral nucleic acid further comprises a photoreceptor-specific

CC promoter. (II) includes photoreceptor promoters providing a means not

CC only for specific targeting of expression in these cells, but also for

CC photoreceptor-restricted transgene expression. This sequence represents a

CC TPL (tripartite leader sequence) from the adenovirus type 5 genome, used

CC to enhance the expression of complementing adenoviral proteins.

XX

SQ Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 1240; DB 10; Length 1240;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATCGCTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60

DB 1 GGATCCACTCTCTCCGATCGCTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60

QY 61 CTGAAAGCGGCGCATGACTTCTCGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120

DB 61 CTGAAAGCGGCGCATGACTTCTCGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120

QY 121 GATATTACATGCGCGCGGATGATCGCTTTGAGGGTGGCGGATCCATCTGGTCAGAAAA 180

DB 121 GATATTACATGCGCGCGGATGATCGCTTTGAGGGTGGCGGATCCATCTGGTCAGAAAA 180

QY 181 GACAATCTTTTGTGTCAAGCTTCGTGGCAAAACGACCGTAGAGGGGTTGGACAGCAA 240

DB 181 GACAATCTTTTGTGTCAAGCTTCGTGGCAAAACGACCGTAGAGGGGTTGGACAGCAA 240

QY 241 CTTGGCGATGGAGCGCAGGGTTGGTTTGTGCGGATCGGCGCGCTCTTTGGCGCGAT 300

DB 241 CTTGGCGATGGAGCGCAGGGTTGGTTTGTGCGGATCGGCGCGCTCTTTGGCGCGAT 300

## RESULT 4

ADF48742

ID ADF48742 standard; DNA; 1240 BP.

XX ADF48742;

XX AC

DT 12-FEB-2004 (first entry)

XX

QY 301 GTTTAGCTGCACGCTATTTCGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGTCGCTC 360

DB 301 GTTTAGCTGCACGCTATTTCGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGTCGCTC 360

QY 361 GTCCGGGACCAAGGTGACGCGCGCAACCGCGCTTGTGAGGGTGCACAGGTCAACGCTGGT 420

DB 361 GTCCGGGACCAAGGTGACGCGCGCAACCGCGCTTGTGAGGGTGCACAGGTCAACGCTGGT 420

QY 421 GGCTACCTCTCCGCGTAGGCGCTGCTGTTGGTCCAGCAGAGCGCGCGCTTCGCGCGAGCA 480

DB 421 GGCTACCTCTCCGCGTAGGCGCTGCTGTTGGTCCAGCAGAGCGCGCGCTTCGCGCGAGCA 480

QY 481 GAATGGCGGTAGGGGTCTAGCTGCGTCTGTCGGGGGGTCTGCGTCCACGGTAAAGAC 540

DB 481 GAATGGCGGTAGGGGTCTAGCTGCGTCTGTCGGGGGGTCTGCGTCCACGGTAAAGAC 540

QY 541 CCCGGGACAGAGCGCGCGCTGCAAGTAGTCTATCTTGCATCCTTTCGAAGTCTAGCGCTG 600

DB 541 CCCGGGACAGAGCGCGCGCTGCAAGTAGTCTATCTTGCATCCTTTCGAAGTCTAGCGCTG 600

QY 601 CTGCCATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGGACCCCATGGCAT 660

DB 601 CTGCCATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGGACCCCATGGCAT 660

QY 661 GGGTGGGTGAGCGCGGAGCGGTACATCCCGCAATCTGCTAAACGTAGAGGGGCTCTCT 720

DB 661 GGGTGGGTGAGCGCGGAGCGGTACATCCCGCAATCTGCTAAACGTAGAGGGGCTCTCT 720

QY 721 GAGTATTCCAAGATATGTAGGCTAGCATCTTCCACCGCGATGCTGGCGCGCATTAATC 780

DB 721 GAGTATTCCAAGATATGTAGGCTAGCATCTTCCACCGCGATGCTGGCGCGCATTAATC 780

QY 781 GTATAGTTCGTGAGGAGCGGAGGAGGTGGGACCGAGGTTCGTACGGGCGGCTGCTC 840

DB 781 GTATAGTTCGTGAGGAGCGGAGGAGGTGGGACCGAGGTTCGTACGGGCGGCTGCTC 840

QY 841 TGCTCGAAGACTATCTGCGTGAAGTGGATGAGTGGATGATATGTTGGACCGCTG 900

DB 841 TGCTCGAAGACTATCTGCGTGAAGTGGATGAGTGGATGATATGTTGGACCGCTG 900

QY 901 GAAGACGTTGAAGCTGCGCTCTGTGAGACCTTACCGCGTACGCAAGAGGAGGCTAGGA 960

DB 901 GAAGACGTTGAAGCTGCGCTCTGTGAGACCTTACCGCGTACGCAAGAGGAGGCTAGGA 960

QY 961 GTCGCGCAGCTTGTGTAACAGCTGCGGCTGACCTGCAAGTTCAGGGCGCAGTAGTCCAG 1020

DB 961 GTCGCGCAGCTTGTGTAACAGCTGCGGCTGACCTGCAAGTTCAGGGCGCAGTAGTCCAG 1020

QY 1021 GGTTCCTTCGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGTTGAG 1080

DB 1021 GGTTCCTTCGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGTTGAG 1080

QY 1081 GACAAACTCTTTCGCGGCTTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGACG 1140

DB 1081 GACAAACTCTTTCGCGGCTTTTCCAGTACTCTTGGATCGGAAACCGCTCCGACG 1140

QY 1141 AGATCCGTACTTCCCGCGCGAGGAGACCTGAGCGAGTCCGCATTCGACCGGATCGGAAACC 1200

DB 1141 AGATCCGTACTTCCCGCGCGAGGAGACCTGAGCGAGTCCGCATTCGACCGGATCGGAAACC 1200

QY 1201 TCTCGAAGAGGCGTCTAAACAGTCCAGTCCAGATCT 1240

DB 1201 TCTCGAAGAGGCGTCTAAACAGTCCAGTCCAGATCT 1240

DE	Ad5 tripartite leader sequence.		
XX			
KW	cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;		
KW	HIV gene expression activation; adenovirus tripartite leader; TPL;		
KW	gutless adenoviral vector particle;		
KW	helper-independent fiberless recombinant adenovirus vector;		
KW	packaging cell line; pseudotyping; adenovirus vector; gene therapy;		
KW	hereditary disorder; tumour; HIV infection; tripartite leader sequence;		
KW	TPL; ds.		
XX			
OS	Human adenovirus type 5.		
XX			
PN	US2003157688-A1.		
XX			
PD	21-AUG-2003.		
XX			
XX	14-JAN-2000; 2000US-00482682.		
XX			
PR	14-JAN-1999; 99US-0115920P.		
PR	26-JUN-2000; 2000US-00423783.		
XX			
PA	(VSEG/) VON SEGGERN D J.		
PA	(NEME/) NEMEROW G R.		
PA	(HALL/) HALLENBECK P.		
PA	(STEV/) STEVENSON S.		
PA	(SKRI/) SKRIPCHENKO Y.		
XX			
XX	Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;		
PI	Skripchenko Y;		
PI			
XX			
DR	WPI; 2003-843463/78.		
XX			
PT	Novel isolated nucleic acid molecule useful for delivering heterologous		
PT	gene to human or any animal, or for producing gutless adenoviral vector		
PT	particle.		
XX			
PS	Claim 14; SEQ ID NO 32; 157pp; English.		
XX			
CC	The invention describes an isolated nucleic acid molecule (I) comprising		
CC	an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide		
CC	sequence comprising a first and second different TPL exons or first,		
CC	second and third same or different TPL exons, the TPL exons chosen from		
CC	complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon		
CC	3. (I) is useful for delivering a heterologous gene to a human or any		
CC	animal, or for producing a gutless adenoviral vector particle. A		
CC	recombinant adenovirus particle (II) is useful for delivery of an		
CC	exogenous gene to a target cell which involves contacting the cell with		
CC	an amount of (II) sufficient to infect the cell. A helper-independent		
CC	fiberless recombinant adenovirus vector genome (III) is useful for		
CC	producing an adenovirus vector particle containing (III) which involves		
CC	providing a packaging cell line which complements replication and		
CC	packaging of the genome and (III) which is deficient in expressing		
CC	sufficient functional fiber protein to support assembly of fiber		
CC	containing particles and harvesting the particle produced by the cell		
CC	line. (III) is useful for pseudotyping recombinant viral vectors which		
CC	involves complementing a missing fiber gene of (III) or helper-dependent		
CC	fiberless recombinant adenovirus vector genome by expressing in packaging		
CC	cells a fiber gene from a different adenoviral serotype than the		
CC	recombinant adenovirus vector. (III) is also useful for specifically		
CC	targeting an adenovirus vector to a cell of choice. (I) is useful for		
CC	gene therapy. (II) is useful for treating diseases such as hereditary		
CC	disorder, and for reducing proliferation of tumour cells in a subject, or		
CC	to disrupt HIV infection. This sequence represents the complete		
CC	Adenovirus tripartite leader sequence.		
XX			
SQ	Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;		
	Query Match 100.0%; Score 1240; DB 10; Length 1240;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 GGATCCACTCTCTTCGCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 60		

Db	1 GGATCCACTCTCTTCGCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 60
QY	61 CTGAAAAGCGGGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
Db	61 CTGAAAAGCGGGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
QY	121 GATATTACCTGCGCGCGGTGATGCCCTTGAAGGTGGCGGATCATCTGTGTGAGAAA 180
Db	121 GATATTACCTGCGCGCGGTGATGCCCTTGAAGGTGGCGGATCATCTGTGTGAGAAA 180
QY	181 GACAATCTTTTGTGTCAGCTTGTGTCGCAAAACCCAGCCGTAGAGGGCGTTCGACAGCAA 240
Db	181 GACAATCTTTTGTGTCAGCTTGTGTCGCAAAACCCAGCCGTAGAGGGCGTTCGACAGCAA 240
QY	241 CTTGCGGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCCGCGAT 300
Db	241 CTTGCGGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCCGCGAT 300
QY	301 GTTTAGCTGCAAGTATTCGCGGCAACCGACCGCCATTTCGGGAAAGACGGTGGTGGCTC 360
Db	301 GTTTAGCTGCAAGTATTCGCGGCAACCGACCGCCATTTCGGGAAAGACGGTGGTGGCTC 360
QY	361 GTGCGGACACAGGTGTCAGCGCCCAACCGGGTTGTGCAAGGTGCAAGGTCAACCTGGT 420
Db	361 GTGCGGACACAGGTGTCAGCGCCCAACCGGGTTGTGCAAGGTGCAAGGTCAACCTGGT 420
QY	421 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGGGCGCGCCCTTCGCGGAGCA 480
Db	421 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGGGCGCGCCCTTCGCGGAGCA 480
QY	481 GAATGGCGGTAGGGGTCTAGCTGGTCTGGTCCGGGGGTCTGGTCCACGGTAAAGAC 540
Db	481 GAATGGCGGTAGGGGTCTAGCTGGTCTGGTCCGGGGGTCTGGTCCACGGTAAAGAC 540
QY	541 CCGCGGACAGCGCGCGCTCGAAGTACTCTATCTTGCATCTTTCGAAGTCTAGCGCTG 600
Db	541 CCGCGGACAGCGCGCGCTCGAAGTACTCTATCTTGCATCTTTCGAAGTCTAGCGCTG 600
QY	601 CTGCGCATCGCGGGCGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660
Db	601 CTGCGCATCGCGGGCGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660
QY	661 GGGTGGGTGAGCGCGGAGGCGTACATCCCGCAATGTCTGTAACGTAAGAGGGGTCTCT 720
Db	661 GGGTGGGTGAGCGCGGAGGCGTACATCCCGCAATGTCTGTAACGTAAGAGGGGTCTCT 720
QY	721 GAGTATTCGAAGATATGATGGTAGCTATCTCCACCGGGATGCTGGCGCGCACGTAATC 780
Db	721 GAGTATTCGAAGATATGATGGTAGCTATCTCCACCGGGATGCTGGCGCGCACGTAATC 780
QY	781 GTATAGTTCTGTCGAGGGAGCGAGGAGTTCGGGACCGAGGTTGCTACGGGGGCTGCTC 840
Db	781 GTATAGTTCTGTCGAGGGAGCGAGGAGTTCGGGACCGAGGTTGCTACGGGGGCTGCTC 840
QY	841 TGCTCGGAAGACTATCTCCCTCAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900
Db	841 TGCTCGGAAGACTATCTCCCTCAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900
QY	901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTACGACGAGAGGCGGTAGGA 960
Db	901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTACGACGAGAGGCGGTAGGA 960
QY	961 GTGCGGACGTTGTGACCGCTGGCGGTGACCTGACCTGACCTGAGGGCGAGTACTCCAG 1020
Db	961 GTGCGGACGTTGTGACCGCTGGCGGTGACCTGACCTGAGGGCGAGTACTCCAG 1020
QY	1021 GGTTCCTTGTGATGATGATCTTATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db	1021 GGTTCCTTGTGATGATGATCTTATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
QY	1081 GACAACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGACG 1140
Db	1081 GACAACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGACG 1140



Db 2049 TCTCGAGAAAGCGTCTTAACCAAGTCACAGTCGCAAGATC 2087

## RESULT 6

ABA94286  
ID ABA94286 standard; DNA; 7231 BP.

XX ABA94286;  
AC ABA94286;

XX 13-MAR-2002 (first entry)

XX Nucleotide sequence of expression plasmid pDV80.

XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;  
KW opthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; fiber protein; ss.

XX Synthetic.

XX WO200183729-A2.

XX 08-NOV-2001.

XX 30-APR-2001; 2001WO-EP004863.

XX 01-MAY-2000; 2000US-00562934.

XX (NOVS ) NOVARTIS AG.

XX (SCRI ) SCRIPPS RES INST.

XX (NEME//) NEMEROW G R.

XX (VSEG//) VON SEGGERN D J.

XX (FRIE//) FRIEDLANDER M.

PI Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2002-082846/11.

XX Polynucleotide for making vectors, useful for treating ocular diseases,  
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat  
PT sequences, packaging signal and photoreceptor-specific promoter.

XX Example 8; Page 146-148; 149pp; English.

XX The invention provides an isolated polynucleotide comprising adenovirus  
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal  
CC operatively linked to ITRS and a photoreceptor-specific promoter. A  
CC recombinant AV vector (AVV) comprising the polynucleotide is useful for  
CC targeted delivery of a gene product to the eye (especially to the  
CC vitreous cavity), for treating an ocular disease, e.g., retinal  
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic  
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal  
CC preferably human. The AVV comprises a fiber protein that specifically or  
CC selectively binds to receptors that are expressed on cells (preferably  
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a  
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein  
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5  
CC penton, and the therapeutic product is a trophic factor, an anti-  
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type  
CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that  
CC regulates expression of a photoreceptor specific gene product. The viral  
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV  
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful  
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber  
CC protein or its portion, and selectively transduces photoreceptors and  
CC delivers a gene product encoded by AAV. The present sequence represents  
CC the nucleotide sequence of plasmid pDV80, an expression plasmid for  
XX adenoviral 37 fiber protein

SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;

Query Match 99.9%;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTTCCGCATCGTGTCTGCGAGGGCCAGCTGTTGGGGTGAAGTACTCCCT 60  
DB |||||  
QY 849 GGATCCACTCTCTTCCGCATCGTGTCTGCGAGGGCCAGCTGTTGGGGTGAAGTACTCCCT 908  
DB |||||  
QY 61 CTGAAAGCGGGCATGACTTCTCGGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTT 120  
DB |||||  
QY 909 CTGAAAGCGGGCATGACTTCTCGGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTT 968  
DB |||||  
QY 121 GATATTACCTGGCCCGCGGTGATGCTTTCAGAGGTGCGCCATCCATCTCTGTCAGAAAA 180  
DB |||||  
QY 969 GATATTACCTGGCCCGCGGTGATGCTTTCAGAGGTGCGCCATCCATCTCTGTCAGAAAA 1028  
DB |||||  
QY 181 GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGCAAGCAAA 240  
DB |||||  
QY 1029 GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGCAAGCAAA 1088  
DB |||||  
QY 241 CTTGGCGATGAGCGCAGGGTTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGCAT 300  
DB |||||  
QY 1089 CTTGGCGATGAGCGCAGGGTTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGCAT 1148  
DB |||||  
QY 301 GTTTAGCTGCACGTATTTCGCGGCAACGACCGCCATTCGCGAAAGACGTTGTCGCTC 360  
DB |||||  
QY 1149 GTTTAGCTGCACGTATTTCGCGGCAACGACCGCCATTCGCGAAAGACGTTGTCGCTC 1208  
DB |||||  
QY 361 GTCGGGCAACAGGTGACGCGCCAAACGCGGTTGTGAGGGTGCACAGGTCAACGCTCGT 420  
DB |||||  
QY 1209 GTCGGGCAACAGGTGACGCGCCAAACGCGGTTGTGAGGGTGCACAGGTCAACGCTCGT 1268  
DB |||||  
QY 421 GGTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTCGCGCAGCA 480  
DB |||||  
QY 1269 GGTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTCGCGCAGCA 1328  
DB |||||  
QY 481 GAATGGCGGTAGGGGTCTAGTCTGCTCTCTCGGGGGTCTCGTCCAGCGTAAAGAC 540  
DB |||||  
QY 1329 GAATGGCGGTAGGGGTCTAGTCTGCTCTCTCGGGGGTCTCGTCCAGCGTAAAGAC 1388  
DB |||||  
QY 541 CCCGGGCAAGCGCGCTCGAAGTAGTCTATCTTTCATCTTTCGCAAGTCTAGGCGCTG 600  
DB |||||  
QY 1389 CCCGGGCAAGCGCGCTCGAAGTAGTCTATCTTTCATCTTTCGCAAGTCTAGGCGCTG 1448  
DB |||||  
QY 601 CTGCCATCTCGCGCGGCAAGCGCGCTCGTATGGTGTGAGTGGGGACCCCATGGCAT 660  
DB |||||  
QY 1449 CTGCCATCTCGCGCGGCAAGCGCGCTCGTATGGTGTGAGTGGGGACCCCATGGCAT 1508  
DB |||||  
QY 661 GGGTGGGTGAGCGCGGAGGCTACATCCGCCAAATGTCGTAAGTAGAGGGGCTCTCT 720  
DB |||||  
QY 1509 GGGTGGGTGAGCGCGGAGGCTACATCCGCCAAATGTCGTAAGTAGAGGGGCTCTCT 1568  
DB |||||  
QY 721 GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCATATC 780  
DB |||||  
QY 1569 GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCATATC 1628  
DB |||||  
QY 781 GTATAGTTCGTGCGAGGGAGCGAGGAGTCCGGACCCAGGTTGCTACGGGGGGTGTCTC 840  
DB |||||  
QY 1629 GTATAGTTCGTGCGAGGGAGCGAGGAGTCCGGACCCAGGTTGCTACGGGGGGTGTCTC 1688  
DB |||||  
QY 841 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTTGGACCGTG 900  
DB |||||  
QY 1689 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTTGGACCGTG 1748  
DB |||||  
QY 901 GAACAGCTTGAAGCTGGCTGTGAGACCTTACCGCTCACGCAAGAGAGGCGTAGGA 960  
DB |||||  
QY 1749 GAACAGCTTGAAGCTGGCTGTGAGACCTTACCGCTCACGCAAGAGAGGCGTAGGA 1808  
DB |||||  
QY 961 GTCGCGCAGCTTGTGTTGACCGAGCTCGGCGGTGACCTGACCGCTCTAGGGCGCAGTAGTCAG 1020  
DB |||||  
QY 1809 GTCGCGCAGCTTGTGTTGACCGAGCTCGGCGGTGACCTGACCGCTCTAGGGCGCAGTAGTCAG 1868  
DB |||||  
QY 1021 GGTTTCTTGTGATGATCATATCTTATCTGTCCTTTTTTTTTTCCACAGCTCGCGGTTGAG 1080  
DB |||||  
QY 1869 GGTTTCTTGTGATGATCATATCTTATCTGTCCTTTTTTTTTTCCACAGCTCGCGGTTGAG 1928  
DB |||||



QY 1081 GACAAACTCTTCGGCGTCTTTCCAGTACTCTTGGATCGGAACACCGTTCGGCTCGGAACG 1140  
Db 1929 GACAAACTCTTCGGCGTCTTTCCAGTACTCTTGGATCGGAACACCGTTCGGCTCGGAACG 1988  
QY 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAACAC 1200  
Db 1989 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAACAC 2048  
QY 1201 TCTCGAGAAAGCGTCTAACCAAGTCACAGTCGCAAGATC 1239  
Db 2049 TCTCGAGAAAGCGTCTAACCAAGTCACAGTCGCAAGATC 2087  
RESULT 7  
ID ADB75132  
XX ADB75132 standard; DNA; 7231 BP.  
XX  
AC ADB75132;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Plasmid pDV80 DNA sequence.  
XX  
KW ophthalmological; antiinflammatory; antidiabetic; gene therapy;  
KW adenovirus inverted terminal repeat sequence;  
KW adenovirus packaging signal; photoreceptor-specific promoter;  
KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;  
KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;  
KW rhodopsin; wild-type Stargardt disease gene; STDG1; anti-cancer agent;  
KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;  
KW diabetic retinopathy; retinal vascularisation; choroideraemia;  
KW gyrate atrophy; macular dystrophy; retinoblastoma;  
KW photoreceptor-restricted transgene expression;  
KW recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;  
KW circular; ds; pDV80.  
XX  
OS Synthetic.  
XX  
PN US2002193327-A1.  
XX  
PD 19-DEC-2002.  
XX  
PF 01-MAY-2001; 2001US-00847101.  
XX  
PR 01-MAY-2000; 2000US-00562934.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Nemerow GR, Von Seggern DJ, Friedlander M;  
XX  
DR WPI; 2003-657234/62.  
XX  
PT Novel nucleic acids comprising adenovirus inverted terminal repeat  
PT sequences, adenovirus packaging signals operatively linked to the  
PT sequences and photoreceptor-specific promoters, useful for treating  
PT retinitis pigmentosa.  
XX  
PS Example 8; Page 99-103; 106pp; English.  
XX  
CC The invention describes an isolated nucleic acid (I) comprising  
CC adenovirus inverted terminal repeat sequence, an adenovirus packaging  
CC signal operatively linked to the sequence, and a photoreceptor-specific  
CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful  
CC for targeted delivery of a gene product to the eye of a mammal which  
CC involves administering (II) that comprises heterologous DNA encoding the  
CC gene product or resulting in expression of the gene product, where the  
CC recombinant virus comprises a fibre protein that specifically or  
CC selectively binds to receptors that are expressed on cells which are  
CC photoreceptors, in the eye. The recombinant virus comprises a fibre  
CC protein which is an adenovirus type 37, from an adenovirus type D  
CC serotype. The fibre is a chimeric protein containing a sufficient portion  
CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for

CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient  
CC portion of an adenovirus serotype D knob portion of the fiber for  
CC selective binding to photoreceptors in the eye of a mammal. The  
CC encapsulated nucleic acid comprises a photoreceptor-specific promoter  
CC operatively linked to a nucleic acid comprising the therapeutic product  
CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding  
CC a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-  
CC cancer agent and a protein that regulates expression of a photoreceptor-  
CC specific gene product. The delivery is effected for treatment of an  
CC ocular disease such as retinal degenerative disease e.g., retinitis  
CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal  
CC vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or  
CC retinoblastoma inherited and acquired retinal and neovascular  
CC degenerative diseases. The viral nucleic acid comprises an adenovirus  
CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging  
CC signal operatively linked to the sequence. The ITRs and packaging signal  
CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or  
CC 5. The viral nucleic acid further comprises a photoreceptor-specific  
CC promoter. (II) includes photoreceptor promoters providing a means not  
CC only for specific targeting of expression in these cells, but also for  
CC photoreceptor-restricted transgene expression. This sequence represents  
CC an adenovirus 37 fibre-expressing plasmid used in the preparation of  
CC adenoviral gene delivery vectors.  
XX

SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;

Query Match 99.9%; Score 1239; DB 10; Length 7231;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGATCCACTCTCTCCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 60  
Db 849 GGATCCACTCTCTCTCCGATCGCTGTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 908  
QY 61 CTGAAAAGCGGGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACGAGGAGGATTT 120  
Db 909 CTGAAAAGCGGGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACGAGGAGGATTT 968  
QY 121 GATATTCACTGGCCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTCGTCAGAAAA 180  
Db 969 GATATTCACTGGCCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTCGTCAGAAAA 1028  
QY 181 GACAATCTTTTGTGTCAGCTTTGGTGGGCAACGACCCGTAGAGGGGTTGGACAGCAA 240  
Db 1029 GACAATCTTTTGTGTCAGCTTTGGTGGGCAACGACCCGTAGAGGGGTTGGACAGCAA 1088  
QY 241 CTTGCGGATGAGCGCAGGGTTTGGTTTTCGCGATCGCGCGCTCTTTGGCGCGCAT 300  
Db 1089 CTTGCGGATGAGCGCAGGGTTTGGTTTTCGCGATCGCGCGCTCTTTGGCGCGCAT 1148  
QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGATTCGGGAAAGACGGTGTGCGCTC 360  
Db 1149 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGATTCGGGAAAGACGGTGTGCGCTC 1208  
QY 361 GTCGGGACACAGGTGACGGCCAAACCGCGTGTGTGCGAGGTGACAGGTCAACGCTGGT 420  
Db 1209 GTCGGGACACAGGTGACGGCCAAACCGCGTGTGTGCGAGGTGACAGGTCAACGCTGGT 1268  
QY 421 GGCTACCTCTCCGCTAGGCGCTGTTGGTCCACGAGGGCGCGCCCTTTGGCGGACGA 480  
Db 1269 GGCTACCTCTCCGCTAGGCGCTGTTGGTCCACGAGGGCGCGCCCTTTGGCGGACGA 1328  
QY 481 GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCGGGGGGGTCTGCGTCCACGGTAAAGAC 540  
Db 1329 GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCGGGGGGGTCTGCGTCCACGGTAAAGAC 1388  
QY 541 CCCGGGACGACGGCGCGGTGAAAGTAGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 600  
Db 1389 CCCGGGACGACGGCGCGGTGAAAGTAGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 1448  
QY 601 CTGCGATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 660  
Db 1449 CTGCGATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 1508



```
QY 661 GGGTGGGTGAGCGCGGAGCGTATCGCCGAAATGTCGTAACGTAGAGGGCTCTCT 720
D 1509 GGGTGGGTGAGCGCGGAGCGTATCGCCGAAATGTCGTAACGTAGAGGGCTCTCT 1568
QY 721 GAGTATCCAGATATGATAGGTAGCATCTTCACCGGAGTCTGGCGCGCACGTAATC 780
D 1569 GAGTATCCAGATATGATAGGTAGCATCTTCACCGGAGTCTGGCGCGCACGTAATC 1628
QY 781 GTATAGTTCTGTCGAGGAGCGAGGAGGTGCGGACCGAGTTGCTACCGGCGGCTGCTC 840
D 1629 GTATAGTTCTGTCGAGGAGCGAGGAGGTGCGGACCGAGTTGCTACCGGCGGCTGCTC 1688
QY 841 TGCTCGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 900
D 1689 TGCTCGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 1748
QY 901 GAAGACGTTGAGCTGGGCTCTGTGAGACCTACCGGTCACGCAAGAGGCGTAGGA 960
D 1749 GAAGACGTTGAGCTGGGCTCTGTGAGACCTACCGGTCACGCAAGAGGCGTAGGA 1808
QY 961 GTCGCGCAGCTGTTGACCAAGCTCGCGGTGACCTGACGCTAGGCGCAGTAGTCCAG 1020
D 1809 GTCGCGCAGCTGTTGACCAAGCTCGCGGTGACCTGACGCTAGGCGCAGTAGTCCAG 1868
QY 1021 GGTTCCTTGATGATGTATCATCTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
D 1869 GGTTCCTTGATGATGTATCATCTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1928
QY 1081 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCGCAAG 1140
D 1929 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCGCAAG 1988
QY 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAC 1200
D 1989 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAC 2048
QY 1201 TCTCGAGAAAGCGCTTAACCAAGTCAACAGTCCGCAAGATC 1239
D 2049 TCTCGAGAAAGCGCTTAACCAAGTCAACAGTCCGCAAGATC 2087
```

## RESULT 8

```
ADFA8774
ID ADFA8774 standard; DNA; 7231 BP.
XX
AC ADFA8774;
XX
DT 12-FEB-2004 (first entry)
XX
DE Adenovirus associated plasmid DNA.
XX
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;
KW HIV gene expression activation; adenovirus tripartite leader; TPL;
KW gutless adenoviral vector particle;
KW helper-independent fiberless recombinant adenovirus vector;
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;
KW hereditary disorder; tumour; HIV infection; ds; cyclic; circular.
XX
OS Synthetic.
OS unidentified adenovirus.
XX
PN US2003157688-A1.
XX
PD 21-AUG-2003.
XX
PF 14-JAN-2000; 2000US-00482682.
XX
PR 14-JAN-1999; 99US-0115920P.
PR 26-JUN-2000; 2000US-00423783.
XX
PA (VSEB/) VON SEGGERN D J.
PA (NEME/) NEMEROW G R.
```

```
PA (HALL/) HALLENBECK P.
PA (STEV/) STEVENSON S.
XX (SKRI/) SKRIPCHENKO Y.
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;
PI Skripchenko Y;
XX WPI; 2003-843463/78.
XX
PT Novel isolated nucleic acid molecule useful for delivering heterologous
PT gene to human or any animal, or for producing gutless adenoviral vector
PT particle.
XX
PS Claim 10; SEQ ID NO 64; 157pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) comprising
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide
CC sequence comprising a first and second different TPL exons or first,
CC second and third same or different TPL exons, the TPL exons chosen from
CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon
CC 3. (I) is useful for delivering a heterologous gene to a human or any
CC animal, or for producing a gutless adenoviral vector particle. A
CC recombinant adenovirus particle (II) is useful for delivery of an
CC exogenous gene to a target cell which involves contacting the cell with
CC an amount of (II) sufficient to infect the cell. A helper-independent
CC fiberless recombinant adenovirus vector genome (III) is useful for
CC producing an adenovirus vector particle containing (III) which involves
CC providing a packaging cell line which complements replication and
CC packaging of the genome and (III) which is deficient in expressing
CC sufficient functional fiber protein to support assembly of fiber
CC containing particles and harvesting the particle produced by the cell
CC line. (III) is useful for pseudotyping recombinant viral vectors which
CC involves complementing a missing fiber gene of (III) or helper-dependent
CC fiberless recombinant adenovirus vector genome by expressing in packaging
CC cells a fiber gene from a different adenoviral serotype than the
CC recombinant adenovirus vector. (III) is also useful for specifically
CC targeting an adenovirus vector to a cell of choice. (I) is useful for
CC gene therapy. (II) is useful for treating diseases such as hereditary
CC disorder, and for reducing proliferation of tumour cells in a subject, or
CC to disrupt HIV infection. This sequence represents a plasmid associated
CC with the creation of adenoviral vectors and packaging cell lines.
XX
SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;
Query Match 99.9%; Score:1239; DB 10; Length 7231;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTCTCCGCATCGCTGTCTGCGAGGCGCCAGCTGTGGGGTGAGTACTCCCT 60
D 849 GGATCCACTCTCTCCGCATCGCTGTCTGCGAGGCGCCAGCTGTGGGGTGAGTACTCCCT 908
QY 61 CTGAAAGCGGGCATGACTTTCGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120
D 909 CTGAAAGCGGGCATGACTTTCGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 968
QY 121 GATATTACCTGGCCCGGGTGATGCCCTTTTCAGGGTGGCCGATCCATCTGTGTGAGAAA 180
D 969 GATATTACCTGGCCCGGGTGATGCCCTTTTCAGGGTGGCCGATCCATCTGTGTGAGAAA 1028
QY 181 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
D 1029 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 1088
QY 241 CTTGGCGATGAGCGCAGGGTTGTTTGTGCGGATCGGCGGCGCTCTTGGCGCGCGAT 300
D 1089 CTTGGCGATGAGCGCAGGGTTGTTTGTGCGGATCGGCGGCGCTCTTGGCGCGCGAT 1148
QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGACCGCCATTTCGGAAGAGCGTGGTGGCTC 360
D 1149 GTTTAGCTGCACGTATTTCGCGCGCAACGACCGCCATTTCGGAAGAGCGTGGTGGCTC 1208
QY 361 GTCGGGCACCGAGTGTACGCGCCCAACCGCGGTGTGTGAGGGGTGACAAGGTCAACGCTG 420
```

```
Db 1209 GTCCGGCAGGAGTSCAGCGCCACCGGGTGTGCGAGGTGACAAAGTCAACGCTGGT 1268
Qy 421 GGCTACTCTCCGCTAGCGCTCGTGGTCCAGCAGAGCGCGCCCTTTGCGGAGCA 480
Db 1269 GGTACCTCTCCGCTAGCGCTCGTGGTCCAGCAGAGCGCGCCCTTTGCGGAGCA 1328
Qy 481 GAATGGCGTAGGGGCTTAGCTGGCTCTCGTCCGGGGGGTCTGGTCCAGGTAAAGAC 540
Db 1329 GAATGGCGTAGGGGCTTAGCTGGCTCTCGTCCGGGGGGTCTGGTCCAGGTAAAGAC 1388
Qy 541 CCCGGCAGCAGCGCGCTGAAAGTAGTCTATCTTTGCATCTCTTCAAGTCTAGCGCTG 600
Db 1389 CCCGGCAGCAGCGCGCTGAAAGTAGTCTATCTTTGCATCTCTTCAAGTCTAGCGCTG 1448
Qy 601 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 660
Db 1449 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 1508
Qy 661 GGGGTGGGTAGCGCGGAGGCTATATCGCGCAATGTCTAAAGTAGAGGGCTCTCT 720
Db 1509 GGGGTGGGTAGCGCGGAGGCTATATCGCGCAATGTCTAAAGTAGAGGGGCTCTCT 1568
Qy 721 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGATGCTGGCGGCACGCTAATC 780
Db 1569 GAGTATTCGAATATGTAGGTAGCATCTTCCACCGCGATGCTGGCGGCACGCTAATC 1628
Qy 781 GTATAGTTCTGTCGAGGAGCGAGAGGTTCGGACCCGAGTTGTCTACGGGCGGGTGCTC 840
Db 1629 GTATAGTTCTGTCGAGGAGCGAGAGGTTCGGACCCGAGTTGTCTACGGGCGGGTGCTC 1688
Qy 841 TGCTCGGAAGACTATCTGCTGAAGATGCGATGCGATGATGATGATGATGATGATGATG 900
Db 1689 TGCTCGGAAGACTATCTGCTGAAGATGCGATGCGATGATGATGATGATGATGATGATG 1748
Qy 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCAAGAGGAGGCTAGGA 960
Db 1749 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCAAGAGGAGGCTAGGA 1808
Qy 961 GTCGCGCAGCTTTGTACAGCTCGCGCGTGACCTGCACTGTAGGGCGCAGTAGTCCAG 1020
Db 1809 GTCGCGCAGCTTTGTACAGCTCGCGCGTGACCTGCACTGTAGGGCGCAGTAGTCCAG 1868
Qy 1021 GGTTCCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1869 GGTTCCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1928
Qy 1081 GACAACTCTTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAACG 1140
Db 1929 GACAACTCTTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAACG 1988
Qy 1141 AGATCCGTACTCCCGCGCGGAGGACCTGAGAGTCCGCAATCGACCGGATCGGAAACCC 1200
Db 1989 AGATCCGTACTCCCGCGCGGAGGACCTGAGAGTCCGCAATCGACCGGATCGGAAACCC 2048
Qy 1201 TCTCGAAGAGCGCTCAACAGTCACAGTCGCAAGATC 1239
Db 2049 TCTCGAAGAGCGCTCAACAGTCACAGTCGCAAGATC 2087
```

```
RESULT 9
AAA59072
ID AAA59072 standard; DNA; 7960 BP.
XX AC
XX AAA59072;
XX 07-NOV-2000 (first entry)
XX Nucleotide sequence of plasmid pdV67.
XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
XX ss.
```

```
OS Synthetic.
XX WO200042208-A1.
XX PD 20-JUL-2000.
XX PF 14-JAN-2000; 2000WO-EP000265.
XX PR 14-JAN-1999; 99US-0115920P.
XX PA (NOVS ) NOVARTIS AG.
XX PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX PA (SCRI ) SCRIPPS RES INST.
XX PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
XX PI Skripchenko Y;
XX DR WPI; 2000-476068/41.
XX PT New nucleic acid comprising an adenovirus tripartite leader nucleotide
XX PT for producing high-capacity and targeted vectors for adenovirus-based
XX PT gene therapy.
XX PS Claim 10; Page 184-186; 212pp; English.
XX CC The specification describes a nucleic acid molecule comprising an
XX CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
XX CC comprising two different TPL exons or three same or different TPL exons.
XX CC The nucleic acid is used to produce an adenovirus vector particle,
XX CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
XX CC vectors, target an adenovirus vector to a cell, produce a modified
XX CC adenovirus, deliver a heterologous gene to an animal and produce a
XX CC gutless adenoviral vector particle. The present sequence represents
XX CC plasmid pdV67, which contains a TPL
SQ Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;
Query Match 99.9%; Score 1239; DB 3; Length 7960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATCCACTCTCTCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
Db 929 GGATCCACTCTCTCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 988
Qy 61 CTGAAAAAGCGGCGATGACTTCTGCGCTAAGATTGCTCAAGATTGCTCAAGATTGCTCAAGATTG 120
Db 989 CTGAAAAAGCGGCGATGACTTCTGCGCTAAGATTGCTCAAGATTGCTCAAGATTGCTCAAGATTG 1048
Qy 121 GATATTACCTGGCGCGGCTGATGCTTTGAGGGTGGCGCGCATCTCGTCAGAAAA 180
Db 1049 GATATTACCTGGCGCGGCTGATGCTTTGAGGGTGGCGCGCATCTCGTCAGAAAA 1108
Qy 181 GACAACTCTTTTGTGTCAAGCTTTGGTGGCAAAACGACCCGTAGAGGGGCTTGGACAGCAA 240
Db 1109 GACAACTCTTTTGTGTCAAGCTTTGGTGGCAAAACGACCCGTAGAGGGGCTTGGACAGCAA 1168
Qy 241 CTTGCGCATGAGCGCAGGGTTGGTTTGTGCGCATGCGCGCGCTCTTTGGCGCGCAT 300
Db 1169 CTTGCGCATGAGCGCAGGGTTGGTTTGTGCGCATGCGCGCGCTCTTTGGCGCGCAT 1228
Qy 301 GTTTAGCTGCACGTATTTCGGCGCAACGACCCGCATTCGGGAAAGAGGTTGGTGGCTC 360
Db 1229 GTTTAGCTGCACGTATTTCGGCGCAACGACCCGCATTCGGGAAAGAGGTTGGTGGCTC 1288
Qy 361 GTCGGCACCCAGGTGTCACGCGCAACCGCGGTTGTGAGGGTGAACAAGTCAACGCTGGT 420
Db 1289 GTCGGCACCCAGGTGTCACGCGCAACCGCGGTTGTGAGGGTGAACAAGTCAACGCTGGT 1348
Qy 421 GGCTACTCTCCGCTAGGCGCTCGTGGTCCAGCAGAGGCGCGCCCTTTGCGGAGCA 480
Db 1349 GGCTACTCTCCGCTAGGCGCTCGTGGTCCAGCAGAGGCGCGCCCTTTGCGGAGCA 1408
```

```
QY 481 GAATGCGGTAGGGGCTTAGCTGCTCTCGTCCGGGGGCTCTGCTCAAGTAAAGAC 540
Db 1409 GAATGCGGTAGGGGCTTAGCTGCTCTCGTCCGGGGGCTCTGCTCAAGTAAAGAC 1468
QY 541 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGCACTCTTCAAGTCTAGCGCTG 600
Db 1469 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGGCATCTTGCAGTCTAGCGCTG 1528
QY 601 CTGCCATGCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGAGCCCATGGCAT 660
Db 1529 CTGCCATGCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGAGCCCATGGCAT 1588
QY 661 GGGGTGGGTAGCGCGGAGCGGTACATGCCGGAATGTCTGTAACCTAGAGGGCTCTCT 720
Db 1589 GGGGTGGGTAGCGCGGAGCGGTACATGCCGGAATGTCTGTAACCTAGAGGGCTCTCT 1648
QY 721 GAGTATTCGAAGATGTAGGCTAGCATCTCCACCGCGGATGCTGGCGCGCAGCTAATC 780
Db 1649 GAGTATTCGAAGATGTAGGCTAGCATCTCCACCGCGGATGCTGGCGCGCAGCTAATC 1708
QY 781 GTATAGTTCTGCGAGGAGCGAGGAGTCCGGACCGAGTTGCTACGGCGGGCTGCTC 840
Db 1709 GTATAGTTCTGCGAGGAGCGAGGAGTCCGGACCGAGTTGCTACGGCGGGCTGCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTCGAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTCGAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 1828
QY 901 GAAGACGTTGAGCTGGGCTCTGTGAGACTACCGGTCACGCACGACGAGGAGCGTAGGA 960
Db 1829 GAAGACGTTGAGCTGGGCTCTGTGAGACTACCGGTCACGCACGACGAGGAGCGTAGGA 1888
QY 961 GTGCGGCGAGCTTGTGACAGCTCGCGCGTGCACGCTGTAGGGCGCAGTAGTCCAG 1020
Db 1889 GTGCGGCGAGCTTGTGACAGCTCGCGCGTGCACGCTGTAGGGCGCAGTAGTCCAG 1948
QY 1021 GGTTCCTTGATGATGTATATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1080
Db 1949 GGTTCCTTGATGATGTATATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 2008
QY 1081 GACAACTCTTCGCGCTCTTCCAGTACTCTTGGATCGGAACCGTCCGACG 1140
Db 2009 GACAACTCTTCGCGCTCTTCCAGTACTCTTGGATCGGAACCGTCCGACG 2068
QY 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGCACTCGCAACCGTCCGAAAC 1200
Db 2069 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGCACTCGCAACCGTCCGAAAC 2128
QY 1201 TCTCGAGAAAGCGCTTAACAGTACAGTCCGCAAGATC 1239
Db 2129 TCTCGAGAAAGCGCTTAACAGTACAGTCCGCAAGATC 2167
```

## RESULT 10

ABA94274  
ID ABA94274 standard; DNA; 7960 BP.

AC ABA94274;

XX 13-MAR-2002 (first entry)

DE Nucleotide sequence of adenoviral plasmid pdv67.

KW Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDGL;  
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; tripartite leader; tPL; ss.

OS Synthetic.

XX WO200183729-A2.

XX 08-NOV-2001.

XX 30-APR-2001; 2001WO-EP004863.  
XX 01-MAY-2000; 2000US-00562934.  
XX (NOVS) NOVARTIS AG.  
PA (SCRI) SCRIPPS RES INST.  
PA (NEME/) NEMEROW G R.  
PA (VSEB/) VON SEGGERN D J.  
XX (FRIE/) FRIEDLANDER M.  
PI Nemerow GR, Von Seggern DJ, Friedlander M;  
XX WPI; 2002-082846/11.

Polynucleotide for making vectors, useful for treating ocular diseases, e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat sequences, packaging signal and photoreceptor-specific promoter.

Example 5; Page 134-136; 149pp; English.

The invention provides an isolated polynucleotide comprising adenovirus (AV) inverter terminal repeat sequences (ITRS), AV packaging signal operatively linked to ITRS and a photoreceptor-specific promoter. A recombinant AV vector (AAV) comprising the polynucleotide is useful for targeted delivery of a gene product to the eye (especially to the vitreous cavity), for treating an ocular disease, e.g., retinal degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularizations, and retinoblastoma, of a mammal preferably human. The AAV comprises a fiber protein that specifically or selectively binds to receptors that are expressed on cells (preferably photoreceptors in the eye). Preferably, the recombinant virus comprise a fiber protein from an adenovirus type D subgroup or is a chimeric protein containing a portion of the N-terminus of an adenovirus type 2 or type 5 penton, and the therapeutic product is a trophic factor, an anti-apoptotic factor, a gene encoding a rhodopsin protein, a wild-type stargardt disease gene (STDGL), an anti-cancer agent and a protein that regulates expression of a photoreceptor specific gene product. The viral nucleic acid of AAV comprises ITRS and packaging signal derived from AAV subgroup B or C, especially an AV type 2 or type 5. AAV is also useful for targeted gene therapy, where the vector comprises an AV type 37 fiber protein or its portion, and selectively transduces photoreceptors and delivers a gene product encoded by AAV. The present sequence represents the nucleotide sequence of plasmid pdv67, a plasmid containing adenoviral tripartite leader (tPL)

Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;

Query Match 99.9%; Score 1239; DB 6; Length 7960;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGCATCGTGTCTGCGAGGGCCAGCTGTGGGGTGTAGTACTCCCT 60  
Db 929 GGATCCACTCTCTCCGCATCGTGTCTGCGAGGGCCAGCTGTGGGGTGTAGTACTCCCT 988  
QY 61 CTGAAAGCGGGCATGACTCTCGCTTAAGATTGTTCAGTTTCCAAAACGAGGAGGATTT 120  
Db 989 CTGAAAGCGGGCATGACTCTCGCTTAAGATTGTTCAGTTTCCAAAACGAGGAGGATTT 1048  
QY 121 GATATTACCTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGTCAGAAAA 180  
Db 1049 GATATTACCTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGTCAGAAAA 1108  
QY 181 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240  
Db 1109 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 1168  
QY 241 CTTGCGATGAGCGCAGGGTGTGTTTCTCGGATCGGGCGCTCTTGGCCGCGCAT 300  
Db 1169 CTTGCGATGAGCGCAGGGTGTGTTTCTCGGATCGGGCGCTCTTGGCCGCGCAT 1228  
QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGCAACCGCATTCGGGMAAGACGGTGGTGGGCTC 360

```
Db 1229 GTTATGCTGCAGTATTTCGGCGCAACGACCGCCATTTCGGGAAGACGGTGGTGGCTC 1288
Qy 361 CTCGGGCAACAGTGCACCGCCAAACCGCGTGTGTCAGGGGTGCAAGGTCAACGCTGGT 420
Db 1289 CTCGGGCAACAGTGCACCGCCAAACCGCGTGTGTCAGGGGTGCAAGGTCAACGCTGGT 1348
Qy 421 GGCTACCTCTCCGCTAGGCGCTCGTGGTCCAGCAGAGGGCGCGCCCTTGC CGCGACGA 480
Db 1349 GGCTACCTCTCCGCTAGGCGCTCGTGGTCCAGCAGAGGGCGCGCCCTTGC CGCGACGA 1408
Qy 481 GAATCGCGGTAGGGGCTCTAGCTCGCTCTCGTCCGGGGGCTCTGGTCCACCGGTAAAGAC 540
Db 1409 GAATCGCGGTAGGGGCTCTAGCTCGCTCTGGTCCGGGGGCTCTGGTCCACCGGTAAAGAC 1468
Qy 541 CCCGGGCAAGCGCGCGCTCGAAGTAGTCTATCTTTCATCTTTCGCAAGTCTAGCGCCTG 600
Db 1469 CCCGGGCAAGCGCGCGCTCGAAGTAGTCTATCTTTCATCTTTCGCAAGTCTAGCGCCTG 1528
Qy 601 CTGCCATCGCGGGGGGCAAGCGCGCTCTGATGGTGTGATGGGGACCCCATGGCAT 660
Db 1529 CTGCCATCGCGGGGGGCAAGCGCGCTCTGATGGTGTGATGGGGACCCCATGGCAT 1588
Qy 661 GGGGTGGGTAGCGCGGAGGCTACATCGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 720
Db 1589 GGGGTGGGTAGCGCGGAGGCTACATCGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 1648
Qy 721 GAGTATCCAAAGATATGAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGTAATC 780
Db 1649 GAGTATCCAAAGATATGAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGTAATC 1708
Qy 781 GTATAGTTCGTCGAGGAGCGAGAGGTTCGGACCGAGGTGCTACGGGCGGGTGTCTC 840
Db 1709 GTATAGTTCGTCGAGGAGCGAGAGGTTCGGACCGAGGTGCTACGGGCGGGTGTCTC 1768
Qy 841 TGCTCGGAAGACTATCTGCTCGAAGATGCGATGTCAGTTGGATGATATGTTGGAAGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTCGAAGATGCGATGTCAGTTGGATGATATGTTGGAAGCTG 1828
Qy 901 GAAGACGTTGAGCTGGCGTCTGTGAGACCTACCGCGTCAGCACGAAGAGGCGGTAGGA 960
Db 1829 GAAGACGTTGAGCTGGCGTCTGTGAGACCTACCGCGTCAGCACGAAGAGGCGGTAGGA 1888
Qy 961 GTCGGCGAGCTTGTGACAGCTCGGCGGTGACCTGCGAGTCTAGGGCGCGAGTAGTCAG 1020
Db 1889 GTCGGCGAGCTTGTGACAGCTCGGCGGTGACCTGCGAGTCTAGGGCGCGAGTAGTCAG 1948
Qy 1021 GGTTCCTTGATGATGTATCTATCTGTCCTCTTTTTCACACAGCTCGCGGTTGAG 1080
Db 1949 GGTTCCTTGATGATGTATCTATCTGTCCTCTTTTTCACACAGCTCGCGGTTGAG 2008
Qy 1081 GACAAACTCTTCGCGGCTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 1140
Db 2009 GACAAACTCTTCGCGGCTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 2068
Qy 1141 AGATCCGTAATCGCGCGCGAGGAGCTGAGGAGTTCGCATCGACCGGATCGGAAAC 1200
Db 2069 AGATCCGTAATCGCGCGCGAGGAGCTGAGGAGTTCGCATCGACCGGATCGGAAAC 2128
Qy 1201 TCTCGAGAAAGCGCTTAACCAAGTCACAGTTCGCAAGATC 1239
Db 2129 TCTCGAGAAAGCGCTTAACCAAGTCACAGTTCGCAAGATC 2167
```

RESULT 11

ADB75120

ID ADB75120 standard; DNA; 7960 BP.

XX AC ADB75120;

XX DT 04-DEC-2003 (first entry)

XX DE Plasmid pdv67 DNA sequence.

ophthalmological; antiinflammatory; antidiabetic; gene therapy;  
adenovirus inverted terminal repeat sequence;  
adenovirus packaging signal; photoreceptor-specific promoter;  
adenovirus type 37; adenovirus type D serotype; adenovirus type 2;  
adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;  
rhodopsin; wild-type Stargardt disease gene; STGD1; anti-cancer agent;  
retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;  
diabetic retinopathy; retinal vascularisation; choroideraemia;  
gyrate atrophy; macular dystrophy; retinoblastoma;  
photoreceptor-restricted transgene expression;  
recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;  
circular; ds; pdv67; tripartite leader sequence; TPL.  
Synthetic.

US2002193327-A1.

19-DEC-2002.

01-MAY-2001; 2001US-00847101.

01-MAY-2000; 2000US-00562934.

(SCRI ) SCRIPPS RES INST.

Nemerow GR, Von Seggern DJ, Friedlander M;

WPI; 2003-657234/62.

Novel nucleic acids comprising adenovirus inverted terminal repeat sequences, adenovirus packaging signals operatively linked to the sequences and photoreceptor-specific promoters, useful for treating retinitis pigmentosa.

Example 5; Page 82-86; 106pp; English.

The invention describes an isolated nucleic acid (I) comprising adenovirus inverted terminal repeat sequence, an adenovirus packaging signal operatively linked to the sequence, and a photoreceptor-specific promoter. A Recombinant adenovirus vector (II) comprising (I) is useful for targeted delivery of a gene product to the eye of a mammal which involves administering (II) that comprises heterologous DNA encoding the gene product or resulting in expression of the gene product, where the recombinant virus comprises a fibre protein that specifically or selectively binds to receptors that are expressed on cells which are photoreceptors, in the eye. The recombinant virus comprises a fibre protein which is an adenovirus type 37, from an adenovirus type D serotype. The fibre is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fibre protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient portion of an adenovirus serotype D knob portion of the fibre for selective binding to photoreceptors in the eye of a mammal. The encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a nucleic acid comprising the therapeutic product which is chosen from tropic factor, anti-apoptotic factor, gene encoding a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-specific gene product. The delivery is effected for treatment of an ocular disease such as retinal degenerative disease e.g., retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or retinoblastoma inherited and acquired retinal and neovascular degenerative diseases. The viral nucleic acid comprises an adenovirus inverted terminal repeat (ITR) sequences, and an adenovirus packaging signal operatively linked to the sequence. The ITRs and packaging signal are derived from an adenovirus serotype B or C, or adenovirus type 2 or 5. The viral nucleic acid further comprises a photoreceptor-specific promoter. (II) includes photoreceptor promoters providing a means not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression. This sequence represents an adenovirus fibre-expressing plasmid for complementation of fibre-gene-deleted adenoviruses that also comprises the adenovirus tripartite leader

CC sequence for enhancing the expression of complementing adenoviral  
CC proteins.  
XX  
SQ Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;  
Query Match 99.9%; Score 1239; DB 10; Length 7960;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGATCCACTCTCTCCGATCGCTCTCTGCGAGGCCAGCTCTGCGGGTGAAGTCTCCCT 60  
DB 929 GGATCCACTCTCTCCGATCGCTCTGCGAGGCCAGCTCTGCGGGTGAAGTCTCCCT 988  
QY 61 CTGAAGAGCGGCGATGACTTCTGCGCTAAGATGTTCAGTTTCCAAAAACGAGAGGATTT 120  
DB 989 CTGAAGAGCGGCGATGACTTCTGCGCTAAGATGTTCAGTTTCCAAAAACGAGAGGATTT 1048  
QY 121 GATATTCACTGCGCGCGGCTGATCCCTTTGAGGGTGGCGCATCCATCTGTCAGAAA 180  
DB 1049 GATATTCACTGCGCGCGGCTGATCCCTTTGAGGGTGGCGCATCCATCTGTCAGAAA 1108  
QY 181 GACAACTCTTTTGTGCTCAAGCTTGGTGGCAACGACCGTAGAGGGCTTGGACAGCAA 240  
DB 1109 GACAACTCTTTTGTGCTCAAGCTTGGTGGCAACGACCGTAGAGGGCTTGGACAGCAA 1168  
QY 241 CTGCGCATGAGCGAGCGGTTTGGTTTGTGCGGATCGCGCGCTCTCTGCGCGCAT 300  
DB 1169 CTGCGCATGAGCGAGCGGTTTGGTTTGTGCGGATCGCGCGCTCTCTGCGCGCAT 1228  
QY 301 GTTTAGCTGACGATATTGCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 360  
DB 1229 GTTTAGCTGACGATATTGCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 1288  
QY 361 GTGCGGACCGAGTGCAGCGCCCAACCGCGTTGTCAGGGTGACAAGTCAACCTCGT 420  
DB 1289 GTGCGGACCGAGTGCAGCGCCCAACCGCGTTGTCAGGGTGACAAGTCAACCTCGT 1348  
QY 421 GCTACCTCTCTCGCTAGCGCTCGTTGTGTCAGAGAGCGCGCGCTCTGCGCGAGCA 480  
DB 1349 GCTACCTCTCTCGCTAGCGCTCGTTGTGTCAGAGAGCGCGCGCTCTGCGCGAGCA 1408  
QY 481 GAATGCGGTAGGGGTCTAGCTGCTCTGTCGGGGGGTCTGCTCCACGTTAAAGAC 540  
DB 1409 GAATGCGGTAGGGGTCTAGCTGCTCTGTCGGGGGGTCTGCTCCACGTTAAAGAC 1468  
QY 541 CCCGGGACGAGCGCGCTGCAAGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 600  
DB 1469 CCCGGGACGAGCGCGCTGCAAGTCTATCTTGCATCTCTGCAAGTCTAGCGCTG 1528  
QY 601 CTGCGATGCGGGGCGCAAGCGCGCTCTGATGGTTGAGTGGGGACCCCATGGCAT 660  
DB 1529 CTGCGATGCGGGGCGCAAGCGCGCTCTGATGGTTGAGTGGGGACCCCATGGCAT 1588  
QY 661 GGGGTGGGTAGCGCGGAGGCTATCATCGCAAAATGCTGTAACGTAAGAGGGCTCTCT 720  
DB 1589 GGGGTGGGTAGCGCGGAGGCTATCATCGCAAAATGCTGTAACGTAAGAGGGCTCTCT 1648  
QY 721 GAGTATTCAGATATGATAGGTAGCATCTTCCACCGGATGCTGCGCGGACGTAATC 780  
DB 1649 GAGTATTCAGATATGATAGGTAGCATCTTCCACCGGATGCTGCGCGGACGTAATC 1708  
QY 781 GTATAGTTGCTGCGAGGAGCGAGAGGTGCGGACCGAGGTTGCTACGGGGGGTCTCT 840  
DB 1709 GTATAGTTGCTGCGAGGAGCGAGAGGTGCGGACCGAGGTTGCTACGGGGGGTCTCT 1768  
QY 841 TGCTCGGAGACTATCTGCTCAAGATGGCATGTGAGTTGATGATATGTTGGACGCTG 900  
DB 1769 TGCTCGGAGACTATCTGCTCAAGATGGCATGTGAGTTGATGATATGTTGGACGCTG 1828  
QY 901 GAAGACGTTGAAGCTGGCTCTGTGAGACCTTACCGGTCACGCCAAGAGGAGCGTAGA 960  
DB 1829 GAAGACGTTGAAGCTGGCTCTGTGAGACCTTACCGGTCACGCCAAGAGGAGCGTAGA 1888

QY 961 GTGCGGACGCTGTGTGACACGCTGCGCGGTGACCTGACAGTCTAGGCGCAGTAGTCCAG 1020  
DB 1889 GTGCGGACGCTGTGTGACACGCTGCGCGGTGACCTGACAGTCTAGGCGCAGTAGTCCAG 1948  
QY 1021 GGTTCCTTGTGATGATGATCATCTTATCTGCTCCCTTTTTCACACAGCTCGCGGTGAG 1080  
DB 1949 GGTTCCTTGTGATGATGATCATCTTATCTGCTCCCTTTTTCACACAGCTCGCGGTGAG 2008  
QY 1081 GACAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCCTCCGAAAG 1140  
DB 2009 GACAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCCTCCGAAAG 2068  
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCCGATCGACCGGATCGGAAACCG 1200  
DB 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCCGATCGACCGGATCGGAAACCG 2128  
QY 1201 TCTCGAGAAAGCGCTCTAACCACTACACAGTCCAGATC 1239  
DB 2129 TCTCGAGAAAGCGCTCTAACCACTACACAGTCCAGATC 2167  
RESULT 12  
ADF48754  
ID ADF48754 standard; DNA; 7960 BP.  
XX  
AC ADF48754;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Fibre expressing/tripartite leader sequence plasmid pDV67.  
XX  
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;  
KW HIV gene expression activation; adenovirus tripartite leader; TPL;  
KW gutless adenoviral vector particle;  
KW helper-independent fiberless recombinant adenovirus vector;  
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;  
KW hereditary disorder; tumour; HIV infection; fibre;  
KW fibre-gene-deleted adenoviruses; hygromycin resistance;  
KW tripartite leader sequence; ds; pCLF; pCDN3/fibre.  
XX  
OS Synthetic.  
OS Human adenovirus type 6.  
XX  
PN US2003157688-A1.  
XX  
PD 21-AUG-2003.  
XX  
PF 14-JAN-2000; 2000US-00482682.  
XX  
PR 14-JAN-1999; 99US-0115920P.  
PR 26-JUN-2000; 2000US-00423783.  
XX  
PA (VSEG/) VON SEGGERN D J.  
PA (NEME/) NEMEROW G R.  
PA (HALL/) HALLENBECK P.  
PA (STEV/) STEVENSON S.  
PA (SKRI/) SKRIPCHENKO Y.  
XX  
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;  
PI Skripchenko Y;  
PI  
XX WPI; 2003-843463/78.  
XX  
PT Novel isolated nucleic acid molecule useful for delivering heterologous  
PT gene to human or any animal, or for producing gutless adenoviral vector  
PT particle.  
XX  
PS Claim 10; SEQ ID NO 44; 157pp; English.  
XX  
CC The invention describes an isolated nucleic acid molecule (I) comprising  
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide  
CC sequence comprising a first and second different TPL exons or first,  
CC second and third same or different TPL exons, the TPL exons chosen from

CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon  
CC 3. (I) is useful for delivering a heterologous gene to a human or any  
CC animal, or for producing a gutless adenoviral vector particle. A  
CC recombinant adenovirus particle (II) is useful for delivery of an  
CC exogenous gene to a target cell which involves contacting the cell with  
CC an amount of (II) sufficient to infect the cell. A helper-independent  
CC fiberless recombinant adenovirus vector genome (III) is useful for  
CC producing an adenovirus vector particle containing (III) which involves  
CC providing a packaging cell line which complements replication and  
CC packaging of the genome and (III) which is deficient in expressing  
CC sufficient functional fiber protein to support assembly of fiber  
CC containing particles and harvesting the particle produced by the cell  
CC line. (III) is useful for pseudotyping recombinant viral vectors which  
CC involves complementing a missing fiber gene of (III) or helper-dependent  
CC fiberless recombinant adenovirus vector genome by expressing in packaging  
CC cells a fiber gene from a different adenoviral serotype than the  
CC recombinant adenovirus vector. (III) is also useful for specifically  
CC targeting an adenovirus vector to a cell of choice. (I) is useful for  
CC gene therapy. (II) is useful for treating diseases such as hereditary  
CC disorder, and for reducing proliferation of tumour cells in a subject, or  
CC to disrupt HIV infection. This sequence represents an adenovirus  
CC tripartite leader sequence added to plasmid pCDN3/fibre to create plasmid  
CC pDV67, an adenovirus fibre expressing plasmid for complementation of E4-  
CC gene-deleted adenoviruses.

SQ Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;

Query Match 99.9%; Score 1239; DB 10; Length 7960;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGATCCACTCTTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 60  
DB 929 GGATCCACTCTTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 988  
QY 61 CTGAAAAGCGGCATGACTTTCGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGTTT 120  
DB 989 CTGAAAAGCGGCATGACTTTCGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGTTT 1048  
QY 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 180  
DB 1049 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 1108  
QY 181 GACATCTTTTGTGTCAGCTTGGTGGCAACACCGGTAGAGGGGTGGACAGCAA 240  
DB 1109 GACATCTTTTGTGTCAGCTTGGTGGCAACACCGGTAGAGGGGTGGACAGCAA 1168  
QY 241 CTGGCGATGGAGCGCAGGGTTTGGTTTTTGTGCGGATCGGCGCTCTTGGCGCGAT 300  
DB 1169 CTGGCGATGGAGCGCAGGGTTTGGTTTTTGTGCGGATCGGCGCTCTTGGCGCGAT 1228  
QY 301 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTTCGGGAAAGACGGTGGTGCCTC 360  
DB 1229 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTTCGGGAAAGACGGTGGTGCCTC 1288  
QY 361 GTCGGGCACAGGTGCACCGCCAAACCGGGTTGTGAGGGTGAACAAGTCAACGCTGGT 420  
DB 1289 GTCGGGCACAGGTGCACCGCCAAACCGGGTTGTGAGGGTGAACAAGTCAACGCTGGT 1348  
QY 421 GGTACCTCTCCGCTAGCGCTCTGTTGTCAGAGAGGGCGCGCTTGGCGAGCA 480  
DB 1349 GGTACCTCTCCGCTAGCGCTCTGTTGTCAGAGAGGGCGCGCTTGGCGAGCA 1408  
QY 481 GAATGGCGGTAGGGGTCTAGTGGCTCTCGTCCGGGGGGTCTGGTCCACGGTAAAGAC 540  
DB 1409 GAATGGCGGTAGGGGTCTAGTGGCTCTCGTCCGGGGGGTCTGGTCCACGGTAAAGAC 1468  
QY 541 CCCGGCAGCAGCGCGGTGCAAGTACTTATCTTGTGATCTTCAAGTCTAGGCGCTG 600  
DB 1469 CCCGGCAGCAGCGCGGTGCAAGTACTTATCTTGTGATCTTCAAGTCTAGGCGCTG 1528  
QY 601 CTGCCATGCGCGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660  
|||||

DB 1529 CTGCCATGCGCGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGACCCCATGGCAT 1588  
QY 661 GGGTGGGTGAGCGCGGAGCGGTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCT 720  
DB 1589 GGGTGGGTGAGCGCGGAGCGGTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCT 1648  
QY 721 GAGTATTCGAAGATATCTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCAGTAAATC 780  
DB 1649 GAGTATTCGAAGATATCTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCAGTAAATC 1708  
QY 781 GTATAGTTCGTGCGAGGGAGCGAGAGGTCCGGACCGAGGTTGCTACGGCGGGCTGCTC 840  
DB 1709 GTATAGTTCGTGCGAGGGAGCGAGAGGTCCGGACCGAGGTTGCTACGGCGGGCTGCTC 1768  
QY 841 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900  
DB 1769 TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 1828  
QY 901 GAAGACCTTGAAGCTGGCTGTGAGACCTTACCGCTCAGCACGAGGAGGGCTAGGA 960  
DB 1829 GAAGACCTTGAAGCTGGCTGTGAGACCTTACCGCTCAGCACGAGGAGGGCTAGGA 1888  
QY 961 GTCCGCGAGCTTGTGACCAAGCTCGCGGTGACCTGACACGTCTAGGCGCAGTAGTCCAG 1020  
DB 1889 GTCCGCGAGCTTGTGACCAAGCTCGCGGTGACCTGACACGTCTAGGCGCAGTAGTCCAG 1948  
QY 1021 GGTTCCTTGATGATGATCATCTTATCTGTGCTTCTTTTTCACAGCTCGCGGTGAG 1080  
DB 1949 GGTTCCTTGATGATGATCATCTTATCTGTGCTTCTTTTTCACAGCTCGCGGTGAG 2008  
QY 1081 GACAACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGACG 1140  
DB 2009 GACAACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGACG 2068  
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAACCG 1200  
DB 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAACCG 2128  
QY 1201 TCTCGAAGAGCGCTTAAACCAAGTCAAGTCCAGTCCGAAGATC 1239  
DB 2129 TCTCGAAGAGCGCTTAAACCAAGTCAAGTCCGAAGATC 2167  
RESULT 13  
AAA59075  
ID AAA59075 standard; DNA; 7989 BP.  
XX  
AC AAA59075;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE Nucleotide sequence of plasmid pDV69.  
XX  
KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;  
XX ss.  
XX Synthetic.  
XX  
XX WO200042208-A1.  
XX  
XX 20-JUL-2000.  
XX  
XX 14-JAN-2000; 2000WO-EP000265.  
XX  
XX 14-JAN-1999; 99US-0115920P.  
XX  
XX (NOVS ) NOVARTIS AG  
XX (NOVS ) NOVARTIS-REFINDUNGEN VERW GES MBH.  
XX (SCRI ) SCRIPPS RES INST.  
XX Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;  
PI Skripchenko Y;  
XX

DR WPI; 2000-476068/41.  
XX New nucleic acid comprising an adenovirus tripartite leader nucleotide  
PT for producing high-capacity and targeted vectors for adenovirus-based  
PT gene therapy.  
XX Claim 10; Page 187-190; 212pp; English.  
XX The specification describes a nucleic acid molecule comprising an  
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence  
CC comprising two different TPL exons or three same or different TPL exons.  
CC The nucleic acid is used to produce an adenovirus vector particle,  
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral  
CC vectors, target an adenovirus vector to a cell, produce a modified  
CC adenovirus, deliver a heterologous gene to an animal and produce a  
CC gutless adenoviral vector particle. The present sequence represents  
CC plasmid pdv69, which contains a TPL  
XX  
SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;  
Query Match 99.9%; Score 1239; DB 3; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGATCACTCTCTCCGATCGCTGTCTGCGAGGCCAGCTGTGGGTGAGTACTCCCT 60  
Db |||||  
QY 61 CTGAAAGCGGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120  
Db |||||  
QY 989 CTGAAAGCGGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 1048  
Db |||||  
QY 121 GATATTCACTCGCCCGCGGTGATCGCTTTGAGGGTGGCGGCATCCATCTGTCAGAAAA 180  
Db |||||  
QY 1049 GATATTCACTCGCCCGCGGTGATCGCTTTGAGGGTGGCGGCATCCATCTGTCAGAAAA 1108  
Db |||||  
QY 181 GACAACTTTTGTGTCAAGCTTGTGCGCAACCGCCTAGAGGGCGTTGGACAGCAA 240  
Db |||||  
QY 1109 GACAACTTTTGTGTCAAGCTTGTGCGCAACCGCCTAGAGGGCGTTGGACAGCAA 1168  
Db |||||  
QY 241 CTTGCGGATGAGCGCAGGGTTGGTTTGTGCGGATCGGCGCGCTCTTGGCGCGCAT 300  
Db |||||  
QY 1169 CTTGCGGATGAGCGCAGGGTTGGTTTGTGCGGATCGGCGCGCTCTTGGCGCGCAT 1228  
Db |||||  
QY 301 GTTTAGCTGCAGTATTCGCGCAACCGCCTAGAGGGCGTTGGACAGCAA 360  
Db |||||  
QY 1229 GTTTAGCTGCAGTATTCGCGCAACCGCCTAGAGGGCGTTGGACAGCAA 1288  
Db |||||  
QY 361 GTGCGGCACCAAGGTGCAGCGCCAAACCGCGTGTGTCAGGGTGAACAAGTCAACGCTGGT 420  
Db |||||  
QY 1289 GTGCGGCACCAAGGTGCAGCGCCAAACCGCGTGTGTCAGGGTGAACAAGTCAACGCTGGT 1348  
Db |||||  
QY 421 GGCTACCTCTCCGCTAGCGCTCGTTGTCAGAGAGGGCGCGCTCTTGGCGGAGCA 480  
Db |||||  
QY 1349 GGCTACCTCTCCGCTAGCGCTCGTTGTCAGAGAGGGCGCGCTCTTGGCGGAGCA 1408  
Db |||||  
QY 481 GAATGCGGTAGGGGTCTAGCTGCTGCTCGTCCGGGGTCTGCTCCAGTGAAGAC 540  
Db |||||  
QY 1409 GAATGCGGTAGGGGTCTAGCTGCTGCTCGTCCGGGGTCTGCTCCAGTGAAGAC 1468  
Db |||||  
QY 541 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTTAGCGCTG 600  
Db |||||  
QY 1469 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTTAGCGCTG 1528  
Db |||||  
QY 601 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660  
Db |||||  
QY 1529 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGTTGAGTGGGGACCCCATGGCAT 1588  
Db |||||  
QY 661 GGGGTGGGTAGCGCGGAGGCGTACATGCCGCAATGCTGTAACGCTAGAGGGGTCTCT 720  
Db |||||  
QY 1589 GGGGTGGGTAGCGCGGAGGCGTACATGCCGCAATGCTGTAACGCTAGAGGGGTCTCT 1648  
Db |||||  
QY 721 GAGTATCCAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 780  
Db |||||

Db ||||| 1649 GAGTATCCAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 1708  
QY ||||| 781 GTATAGTTCTGTCGAGGAGGAGGAGGTCCGGACCGAGGTTGCTACCGGGCGGCTGCTC 840  
Db ||||| 1709 GTATAGTTCTGTCGAGGAGGAGGAGGTCCGGACCGAGGTTGCTACCGGGCGGCTGCTC 1768  
QY ||||| 841 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900  
Db ||||| 1769 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 1828  
QY ||||| 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTCAGCAGCAGAGGAGGCTAGGA 960  
Db ||||| 1829 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTCAGCAGCAGAGGAGGCTAGGA 1888  
QY ||||| 961 GTCCGCGAGCTGTGTTGACACGCTCGCGGTGACCTGACGCTCTAGGGCGCAGTACTCCAG 1020  
Db ||||| 1889 GTCCGCGAGCTGTGTTGACACGCTCGCGGTGACCTGACGCTCTAGGGCGCAGTACTCCAG 1948  
QY ||||| 1021 GGTTCCTTGATGATGATCATCTTATCTGCTCCCTTTTTTTCACAGCTCGCGTTGAG 1080  
Db ||||| 1949 GGTTCCTTGATGATGATCATCTTATCTGCTCCCTTTTTTTCACAGCTCGCGTTGAG 2008  
QY ||||| 1081 GACAAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 1140  
Db ||||| 2009 GACAAACTCTTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 2068  
QY ||||| 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAACCG 1200  
Db ||||| 2069 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAACCG 2128  
QY ||||| 1201 TCTCGAGAAAGCGCTCTAAACCACTGACAGTCCGCAAGATC 1239  
Db ||||| 2129 TCTCGAGAAAGCGCTCTAAACCACTGACAGTCCGCAAGATC 2167  
RESULT 14  
ABA94277  
ID ABA94277 standard; DNA; 7989 BP.  
XX ABA94277;  
XX  
XX 13-MAR-2002 (first entry)  
XX Nucleotide sequence of adenoviral plasmid pdv69.  
XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;  
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;  
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;  
KW gene therapy; fiber protein; ss.  
XX Synthetic.  
XX WO200183729-A2.  
XX 08-NOV-2001.  
XX 30-APR-2001; 2001WO-EP004863.  
XX 01-MAY-2000; 2000US-00562934.  
XX (NOVS ) NOVARTIS AG.  
PA (SCRI ) SCRIPPS RES INST.  
PA (NEME/) NEMEROW G R.  
PA (VSEG/) VON SEGGERN D J.  
XX (FRIE/) FRIEDLANDER M.  
PI Nemerow GR, Von Seggern DJ, Friedlander M;  
XX WPI; 2002-082846/11.  
DR Polynucleotide for making vectors, useful for treating ocular diseases,  
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat



PT sequences, packaging signal and photoreceptor-specific promoter.

XX Example 5; Page 137-139; 149pp; English.

PS The invention provides an isolated polynucleotide comprising adenovirus

CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal

CC operatively linked to ITRS and a photoreceptor-specific promoter. A

CC recombinant AV vector (AV) comprising the polynucleotide is useful for

CC targeted delivery of a gene product to the eye (especially to the

CC vitreous cavity), for treating an ocular disease, e.g., retinal

CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic

CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal

CC preferably human. The AAV comprises a fiber protein that specifically or

CC selectively binds to receptors that are expressed on cells (preferably

CC photoreceptors in the eye). Preferably, the recombinant virus comprise a

CC fiber protein from an adenovirus type D subgroup or is a chimeric protein

CC containing a portion of the N-terminus of an adenovirus type 2 or type 5

CC penton, and the therapeutic product is a trophic factor, an anti-

CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type

CC stargardt disease gene (STDB1), an anti-cancer agent and a protein that

CC regulates expression of a photoreceptor specific gene product. The viral

CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV

CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful

CC for targeted gene therapy, where the vector comprises an AV type 37 fiber

CC protein or its portion, and selectively transduces photoreceptors and

CC delivers a gene product encoded by AAV. The present sequence represents

CC the nucleotide sequence of plasmid pdV69, a plasmid containing a modified

XX adenoviral fiber protein

SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;

Query Match 99.9%; Score 1239; DB 6; Length 7989;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATGCTCTCTCGAGGCGCAGCTGTGGGTGAGTACTCCCT 60

DB 929 GGATCCACTCTCTCCGATGCTCTCTCGAGGCGCAGCTGTGGGTGAGTACTCCCT 988

QY 61 CTGAAAAGCGGCGATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120

DB 989 CTGAAAAGCGGCGATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 1048

QY 121 GATATTACCTGGCCGCGGTGATCGCTTTAGAGGTGGCGCATCCATCTGGTCAGAAAA 180

DB 1049 GATATTACCTGGCCGCGGTGATCGCTTTAGAGGTGGCGCATCCATCTGGTCAGAAAA 1108

QY 181 GACAACTCTTTGTGTCAGCTTGTGCGCAACCGCCGATAGAGGGGTTGGACAGCA 240

DB 1109 GACAACTCTTTGTGTCAGCTTGTGCGCAACCGCCGATAGAGGGGTTGGACAGCA 1168

QY 241 CTTGGCGATGAGCGAGCGTTGGTTGTTTTCGCGATCGGCGCTCTTTGGCGCGAT 300

DB 1169 CTTGGCGATGAGCGAGCGTTGGTTGTTTTCGCGATCGGCGCTCTTTGGCGCGAT 1228

QY 301 GTTTAGTGCAGATTTCGCGCGCAACCGCACCGCCATTTCGGGAAAGACGGTGGCGCTC 360

DB 1229 GTTTAGTGCAGATTTCGCGCGCAACCGCACCGCCATTTCGGGAAAGACGGTGGCGCTC 1288

QY 361 CTCGGGACACCGTGCAGCGCCACCGGTTGTGAGGTGACAGGTCAACGCTGGT 420

DB 1289 GTTCGGGACACCGTGCAGCGCCACCGGTTGTGAGGTGACAGGTCAACGCTGGT 1348

QY 421 GCTACCTCTCCGCTAGCGCTCGTGTGTCAGCAGAGGGCGGCGCTTCGCGAGCA 480

DB 1349 GCTACCTCTCCGCTAGCGCTCGTGTGTCAGCAGAGGGCGGCGCTTCGCGAGCA 1408

QY 481 GAATCGCGTAGGGGTCTAGCTGCGTCTGTCGCGGGGGTCTGCGTCCACGGTTAAGAC 540

DB 1409 GAATCGCGTAGGGGTCTAGCTGCGTCTGTCGCGGGGGTCTGCGTCCACGGTTAAGAC 1468

QY 541 CCCGGGACGAGCGCGGTTCGAGTCTATCTTTCATCTTTCGAGTCTAGCGCTG 600

DB 1469 CCCGGGACGAGCGCGCGCTCGAAGTAGTCTATCTTGCATCTTTCGAAGTCTAGCGCCTG 1528

QY 601 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGGGTGAGTGGGGGACCCCATGGCAT 660

DB 1529 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGGGTGAGTGGGGGACCCCATGGCAT 1588

QY 661 GGGGTGGGTGAGCGCGGAGGGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 720

DB 1589 GGGGTGGGTGAGCGCGGAGGGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 1648

QY 721 GAGTATTTCCAAAGATATGTAGGTAGCATCTTCCACCGCGATGTGGCGCCACGTAATC 780

DB 1649 GAGTATTTCCAAAGATATGTAGGTAGCATCTTCCACCGCGATGTGGCGCCACGTAATC 1708

QY 781 GTATAGTTCGTGCGAGGAGCGAGAGGTGGGACCGAGGTTGTACGGGCGGCTGCTC 840

DB 1709 GTATAGTTCGTGCGAGGAGCGAGAGGTGGGACCGAGGTTGTACGGGCGGCTGCTC 1768

QY 841 TGCTCGGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGACCGCTG 900

DB 1769 TGCTCGGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGACCGCTG 1828

QY 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGAGGAGGCGTAGGA 960

DB 1829 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGAGGAGGCGTAGGA 1888

QY 961 GTGCGCAGCTTGTGACAGCTCGGCGGTGACCTGCACTGAGGTGAGGCGCAGTAGTCCAG 1020

DB 1889 GTGCGCAGCTTGTGACAGCTCGGCGGTGACCTGCACTGAGGTGAGGCGCAGTAGTCCAG 1948

QY 1021 GGTTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACAGCTCGCGGTTGAG 1080

DB 1949 GGTTTCCTTGATGATGTCATCTTATCTGCTCCCTTTTTCACAGCTCGCGGTTGAG 2008

QY 1081 GACAACTCTTTCGCGGCTTTTCAGTACTCTTGTGATCGGAAACCGCTCGGCTCCGAAACG 1140

DB 2009 GACAACTCTTTCGCGGCTTTTCAGTACTCTTGTGATCGGAAACCGCTCGGCTCCGAAACG 2068

QY 1141 AGATCCGTACTTCCGCCCGGAGGACCTGACGCGAGTCCGATCGCATCGGATCGGAAAC 1200

DB 2069 AGATCCGTACTTCCGCCCGGAGGACCTGACGCGAGTCCGATCGCATCGGATCGGAAAC 2128

QY 1201 TCTCGAAGAGGCTCTAACCACTGTCAGTCCGCAAGATC 1239

DB 2129 TCTCGAAGAGGCTCTAACCACTGTCAGTCCGCAAGATC 2167

RESULT 15

ADB75123

ID ADB75123 standard; DNA; 7989 BP.

XX AC ADB75123;

XX AC ADB75123;

DT 04-DEC-2003 (first entry)

XX Plasmid pdV69 DNA sequence.

DE ophthalmological; antiinflammatory; antidiabetic; gene therapy;

KW adenovirus inverted terminal repeat sequence;

KW adenovirus packaging signal; photoreceptor-specific promoter;

KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;

KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;

KW rhodopsin; wild-type Stargardt disease gene; STDB1; anti-cancer agent;

KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;

KW diabetic retinopathy; retinal vascularisation; choroidaemia;

KW gyrate atrophy; macular dystrophy; retinoblastoma;

KW photoreceptor-restricted transgene expression;

KW recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;

KW circular; ds; pdV69; tripartite leader sequence; TPL.

OS Synthetic.

XX US2002193327-A1.

PN



XX PD 19-DEC-2002.  
 XX PF 01-MAY-2001; 2001US-00847101.  
 XX PR 01-MAY-2000; 2000US-00562934.  
 XX PA (SRI ) SCRIPPS RES INST.  
 XX PI Nemerow GR, Von Seggern DJ, Friedlander M;  
 XX DR WPI; 2003-657234/62.  
 XX PT Novel nucleic acids comprising adenovirus inverted terminal repeat  
 PT sequences, adenovirus packaging signals operatively linked to the  
 PT sequences and photoreceptor-specific promoters, useful for treating  
 PT retinitis pigmentosa.  
 XX PS Example 5; Page 86-90; 106pp; English.  
 XX CC The invention describes an isolated nucleic acid (I) comprising  
 CC adenovirus inverted terminal repeat sequence, an adenovirus packaging  
 CC signal operatively linked to the sequence, and a photoreceptor-specific  
 CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful  
 CC for targeted delivery of a gene product to the eye of a mammal which  
 CC involves administering (II) that comprises heterologous DNA encoding the  
 CC gene product or resulting in expression of the gene product, where the  
 CC recombinant virus comprises a fibre protein that specifically or  
 CC selectively binds to receptors that are expressed on cells which are  
 CC photoreceptors, in the eye. The recombinant virus comprises a fibre  
 CC protein which is an adenovirus type 37, from an adenovirus type B  
 CC serotype. The fibre is a chimeric protein containing a sufficient portion  
 CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for  
 CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient  
 CC portion of an adenovirus serotype D knob portion of the fiber for  
 CC selective binding to photoreceptors in the eye of a mammal. The  
 CC encapsulated nucleic acid comprises a photoreceptor-specific promoter  
 CC operatively linked to a nucleic acid comprising the therapeutic product  
 CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding  
 CC a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-  
 CC cancer agent and a protein that regulates expression of a photoreceptor-  
 CC specific gene product. The delivery is effected for treatment of an  
 CC ocular disease such as retinal degenerative disease e.g., retinitis  
 CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal  
 CC vascularisation, choroideremia, gyrate atrophy or macular dystrophy or  
 CC retinoblastoma inherited and acquired retinal and neovascular  
 CC degenerative diseases. The viral nucleic acid comprises an adenovirus  
 CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging  
 CC signal operatively linked to the sequence. The ITRs and packaging signal  
 CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or  
 CC 5. The viral nucleic acid further comprises a photoreceptor-specific  
 CC promoter. (II) includes photoreceptor promoters providing a means not  
 CC only for specific targeting of expression in these cells, but also for  
 CC photoreceptor-restricted transgene expression. This sequence represents  
 CC an adenovirus fibre-expressing plasmid for complementation of fibre-gene-  
 CC deleted adenoviruses that also comprises the adenovirus tripartite leader  
 CC sequence for enhancing the expression of complementing adenoviral  
 CC proteins.  
 XX SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;  
 Query Match 99.9%; Score 1239; DB 10; Length 7989;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGATCCACTCTCTCCGATCGCTCTGCGAGGCGCCAGCTGTGGGGTGGAGTACCTCCCT 60  
 DB 929 GGAATCCACTCTCTCCGATCGCTCTGCGAGGCGCCAGCTGTGGGGTGGAGTACCTCCCT 988  
 QY 61 CTGAAGAGGGGCGATGCTTCGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATT 120  
 DB 989 CTGAAGAGGGGCGATGCTTCGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATT 1048

QY 121 GATATTCACTGGCCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 180  
 DB 1049 GATATTCACTGGCCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 1108  
 QY 181 GACAACTCTTTTGTGTCAGCTTGTGCGAAACGACCCGTAGAGGGGTGGACAGCAA 240  
 DB 1109 GACAACTCTTTTGTGTCAGCTTGTGCGAAACGACCCGTAGAGGGGTGGACAGCAA 1169  
 QY 241 CTTGGCGATGAGCGCAGGGTGTGGTTTGTGCGGATCGCGCGGCTCTTTGGCCGCGAT 300  
 DB 1169 CTTGGCGATGAGCGCAGGGTGTGGTTTGTGCGGATCGCGCGGCTCTTTGGCCGCGAT 1228  
 QY 301 GTTTAGTGCACGTATTTCGCCGCAACGACCCGCGATTCGGGAAAGACGGTGTGGCTC 360  
 DB 1229 GTTTAGTGCACGTATTTCGCCGCAACGACCCGCGATTCGGGAAAGACGGTGTGGCTC 1288  
 QY 361 GTCCGGGACCAAGGTGACGCGCCAAACCGCGGTGTGCGAGGGTGACAGGTCAACGCTGGT 420  
 DB 1289 GTCCGGGACCAAGGTGACGCGCCAAACCGCGGTGTGCGAGGGTGACAGGTCAACGCTGGT 1348  
 QY 421 GGCTACCTCTCCGCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTGCCGAGCA 480  
 DB 1349 GGCTACCTCTCCGCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTGCCGAGCA 1408  
 QY 481 GAATGGCGGTAGGGGTCTAGCTCGTCTCGTCCGGGGGTCTCGCTCCACGGTAAAGAC 540  
 DB 1409 GAATGGCGGTAGGGGTCTAGCTCGTCTCGTCCGGGGGTCTCGCTCCACGGTAAAGAC 1468  
 QY 541 CCGGGGACGAGCGCGCGGTGAGTGTCTATCTTCGATCTTTCGAGTCTAGGCGCTG 600  
 DB 1469 CCGGGGACGAGCGCGCGGTGAGTGTCTATCTTCGATCTTTCGAGTCTAGGCGCTG 1528  
 QY 601 CTGCGATCGCGCGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 660  
 DB 1529 CTGCGATCGCGCGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 1588  
 QY 661 GGGTGGGTGAGCGCGAGGCGTACATCGCCAAATGTCGTAAGAGGGGTCTCT 720  
 DB 1589 GGGTGGGTGAGCGCGAGGCGTACATCGCCAAATGTCGTAAGAGGGGTCTCT 1648  
 QY 721 GAGTATTCGAAGATATGAGGTAGCTTCCACCGCGATGCTGGCGCGCACGTATC 780  
 DB 1649 GAGTATTCGAAGATATGAGGTAGCTTCCACCGCGATGCTGGCGCGCACGTATC 1708  
 QY 781 GTATAGTTCGTGCGAGGAGCGAGGTGGGACCGAGGTTGCTACGGCGGGCTGCTC 840  
 DB 1709 GTATAGTTCGTGCGAGGAGCGAGGTGGGACCGAGGTTGCTACGGCGGGCTGCTC 1768  
 QY 841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900  
 DB 1769 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 1828  
 QY 901 GAAGACGTTGAGCTGGCGTGTGAGACCTTACCGGTACCGACGAGAGGGGTAGGA 960  
 DB 1829 GAAGACGTTGAGCTGGCGTGTGAGACCTTACCGGTACCGACGAGAGGGGTAGGA 1888  
 QY 961 GTCCGCGAGCTTGTGACCGAGCTCGCGGTGACCTGCACGCTAGGCGCGAGTGTCCAG 1020  
 DB 1889 GTCCGCGAGCTTGTGACCGAGCTCGCGGTGACCTGCACGCTAGGCGCGAGTGTCCAG 1948  
 QY 1021 GGTTCCTTGTATGATGATCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080  
 DB 1949 GGTTCCTTGTATGATGATCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 2008  
 QY 1081 GACAAACTCTTCGCGGTCTTTCAGTACTCTTTGGATCGGAAACCCCGTGGGCTCCGAACG 1140  
 DB 2009 GACAAACTCTTCGCGGTCTTTCAGTACTCTTTGGATCGGAAACCCCGTGGGCTCCGAACG 2068  
 QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCGGAGTCCGATTCGACCGGATCGGAAACC 1200  
 DB 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGCGGAGTCCGATTCGACCGGATCGGAAACC 2128  
 QY 1201 TCTCGAGAAAGGGCGTCTAAACCAAGTCACAGTCGCAAGATC 1239

Db 2129 TCTCGAGAAAGCGTCTTAACCAAGTCACAGTCACAGTCGCAAGATC 2167  
|||||

Search completed: July 14, 2005, 07:01:37  
Job time : 1751.31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 11806.6 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-32  
Perfect score: 1240  
Sequence: 1 ggatccactctctccgcatt.....cagtcacagtcgcaagatct 1240

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gsl1:\*
- 9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116.2	9.4	617	9	CL706515
2	115.6	9.3	581	9	CL610425
3	115.6	9.3	637	9	CG899744
4	114.6	9.2	404	9	CL266184
5	108.6	8.8	717	9	CL256302
6	103.6	8.4	728	9	CL256297
7	100.2	8.1	640	9	CL706139
8	88	7.1	579	9	CG691637
9	62.2	5.0	925	9	CNS0091P
10	58.2	4.7	925	9	CNS0091P
11	53.6	4.3	645	9	CC721633
12	52.8	4.3	728	2	BE704468
13	49.2	4.0	645	9	CNS01213
14	48.8	3.9	525	4	BM321004
15	48.2	3.9	516	9	CC624407
16	47.4	3.8	716	4	BG873665
17	47.2	3.8	1012	9	CL476462
18	47	3.8	502	6	CA777137
19	47	3.8	537	5	B0777964
20	47	3.8	3475	3	BC036198
21	46.8	3.8	408	9	CL956771
22	46.8	3.8	604	9	CG049742
23	46.8	3.8	605	9	CG049740
24	46.8	3.8	716	6	CA223022

25	46.8	3.8	889	9	CG344984
26	46.8	3.8	2598	3	AY103647
27	46.6	3.8	463	7	CK122370
28	46.6	3.8	585	4	BI776330
29	46.6	3.8	634	7	CV057862
30	46.6	3.8	704	6	CD935513
31	46.6	3.8	721	7	CV054175
32	46.6	3.8	912	5	BX384097
33	46.6	3.8	942	2	BE214157
34	46.4	3.7	763	4	BI157458
35	46.2	3.7	619	6	CA086410
36	46	3.7	619	9	AG114606
37	46	3.7	798	9	CG433993
38	46	3.7	910	9	CNS0060N
39	46	3.7	1201	9	CNS014BJ
40	45.8	3.7	1133	7	CK209664
41	45.8	3.7	1134	5	BM915656
42	45.8	3.7	1947	9	CL970284
43	45.8	3.7	2332	9	AG363333
44	45.6	3.7	541	6	CD931334
45	45.4	3.7	577	6	CD874810

ALIGNMENTS

RESULT 1  
CL706515  
LOCUS FHCRG-GT-S22-5F1 2K-GTA Mus musculus cDNA clone FHCRG-GT-S22-5F 5', linear GSS 26-JUL-2004  
DEFINITION mRNA sequence.  
ACCESSION CL706515  
VERSION CL706515.1 GI:50593553  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 617)  
AUTHORS Soriano,P.  
TITLE www.fhcr.org/labs/soriano/  
JOURNAL Unpublished (2003)  
COMMENT Contact: Soriano P  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' intertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at  
https://www.fhcr.org/labs/soriano/OTdb/  
Class: Gene Trap.  
Location/Qualifiers  
1. 617  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRG-GT-S22-5F"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"  
/clone\_lib="2K-GTA"  
/note="Vector: ROSAFARY"

FEATURES  
source  
1. 617  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRG-GT-S22-5F"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"  
/clone\_lib="2K-GTA"  
/note="Vector: ROSAFARY"

ORIGIN  
Query Match 9.4%; Score 116.2; DB 9; Length 617;  
Best Local Similarity 84.4%; Pred. No. 1.2e-20;  
Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 107  
 |||||  
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 60  
 |||||  
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCATCCA 167  
 |||||  
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCGCA 120  
 |||||  
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201  
 |||||  
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154  
 |||||

RESULT 2  
 CL610425 581 bp mRNA linear GSS 01-JUL-2004  
 LOCUS  
 DEFINITION FHRCR-GT-S17-4D1 2K-GTA Mus musculus cDNA clone FHRCR-GT-S17-4D 5',  
 mRNA sequence.  
 ACCESSION CL610425  
 VERSION  
 KEYWORDS  
 SOURCE GSS.  
 ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Soriano, P.  
 TITLE www.fhrcr.org/labs/soriano/  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Soriano P  
 Division of Basic Sciences, A2-025  
 Fred Hutchinson Cancer Research Center  
 1100 Fairview Ave. N., Seattle, WA 98109, USA  
 Tel: 206 667 6825  
 Fax: 206 667 6522  
 Email: psoriano@fhrcr.org  
 ROSAFARY gene trap. The sequence tag is generated by 3'RACE and  
 represents the 3' insertional cDNA flanking sequence. Additional  
 information regarding this ES cell line and the insertion mutation  
 is available upon request at  
 https://www.fhrcr.org/labs/soriano/Gtdb/  
 Class: Gene Trap.

FEATURES  
 source  
 1..581  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129S4"  
 /db\_xref="taxon:10090"  
 /clone="FHRCR-GT-S17-4D"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="AK7.1"  
 /clone\_lib="2K-GTA"  
 /note="Vector: ROSAFARY"

ORIGIN  
 Query Match 9.3%; Score 115.6; DB 9; Length 581;  
 Best Local Similarity 84.4%; Pred. No. 1.8e-20;  
 Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 107  
 |||||  
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 60  
 |||||  
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCATCCA 167  
 |||||  
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCGCA 120  
 |||||  
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201  
 |||||  
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154  
 |||||

Query Match 9.3%; Score 115.6; DB 9; Length 581;  
 Best Local Similarity 84.4%; Pred. No. 1.8e-20;  
 Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 107  
 |||||  
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 60  
 |||||  
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCATCCA 167  
 |||||  
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCGCA 120  
 |||||  
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201  
 |||||  
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154  
 |||||

RESULT 3  
 CG899744  
 LOCUS  
 DEFINITION FHRCR-GT-S9-7G1 2K-GTA Mus musculus genomic clone FHRCR-GT-S9-7G1  
 5', genomic survey sequence.  
 ACCESSION CG899744  
 VERSION  
 KEYWORDS  
 SOURCE GSS.  
 ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Soriano, P.  
 TITLE www.fhrcr.org/labs/soriano/  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Soriano P  
 Division of Basic Sciences, A2-025  
 Fred Hutchinson Cancer Research Center  
 1100 Fairview Ave. N., Seattle, WA 98109, USA  
 Tel: 206 667 6825  
 Fax: 206 667 6522  
 Email: psoriano@fhrcr.org  
 ROSAFARY gene trap. The sequence tag is generated by 3'RACE and  
 represents the 3' insertional cDNA flanking sequence. Additional  
 information regarding this ES cell line and the insertion mutation  
 is available upon request at  
 https://www.fhrcr.org/labs/soriano/Gtdb/  
 Class: Gene Trap.

FEATURES  
 source  
 1..637  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129S4"  
 /db\_xref="taxon:10090"  
 /clone="FHRCR-GT-S9-7G1"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="AK7.1"  
 /clone\_lib="2K-GTA"  
 /note="Vector: ROSAFARY"

ORIGIN  
 Query Match 9.3%; Score 115.6; DB 9; Length 637;  
 Best Local Similarity 84.4%; Pred. No. 1.8e-20;  
 Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 107  
 |||||  
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 60  
 |||||  
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCATCCA 167  
 |||||  
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCGCA 120  
 |||||  
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201  
 |||||  
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154  
 |||||

Query Match 9.3%; Score 115.6; DB 9; Length 637;  
 Best Local Similarity 84.4%; Pred. No. 1.8e-20;  
 Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 107  
 |||||  
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTCAGTTTCCAAAA 60  
 |||||  
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCATCCA 167  
 |||||  
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGGCCGCGCGCA 120  
 |||||  
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201  
 |||||  
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154  
 |||||

RESULT 4  
 CL266184  
 LOCUS  
 DEFINITION FHRCR-GT-S12-1B1 2K-GTA Mus musculus genomic clone FHRCR-GT-S12-1B1  
 5', genomic survey sequence.  
 ACCESSION CL266184  
 VERSION  
 KEYWORDS  
 SOURCE GSS.  
 ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 404)

AUTHORS Soriano, P.  
TITLE www.fhcr.org/labs/soriano/  
JOURNAL Unpublished (2003)  
COMMENT Contact: Soriano P  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at  
https://www.fhcr.org/labs/soriano/GTdb/  
Class: Gene Trap.

FEATURES source  
Location/Qualifiers  
1..404  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRC-GT-S12-1B1"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"  
/clone\_lib="2K-GTA"  
/note="Vector: ROSAFARY"

ORIGIN  
Query Match 9.2%; Score 114.6; DB 9; Length 404;  
Best Local Similarity 84.3%; Pred. No. 3.1e-20;  
Matches 129; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 49 TGAGTACTCCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAA 108  
Db 1 TGAGTACTCCCTCTCTCAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAA 60  
QY 109 CGAGGAGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCAT 168  
Db 61 CGAGGAGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCAT 120  
QY 169 CTGCTCAGAAAAGACAATCTTTTGTGTCAG 201  
Db 121 CGAAGTCTCTATCCGAAGTTCCTATTCTCTAG 153

RESULT 5  
CL256302  
LOCUS FHCRC-GT-S10-8F1 717 bp DNA linear GSS 10-FEB-2004  
DEFINITION FHCRC-GT-S10-8F1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8F1  
5', genomic survey sequence.  
ACCESSION CL256302  
VERSION CL256302.1 GI:41359955  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 717)  
Contact: Soriano P  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation

AUTHORS Soriano, P.  
TITLE www.fhcr.org/labs/soriano/  
JOURNAL Unpublished (2003)  
COMMENT Contact: Soriano P  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at  
https://www.fhcr.org/labs/soriano/GTdb/  
Class: Gene Trap.

FEATURES source  
Location/Qualifiers  
1..717  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRC-GT-S10-8F1"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"  
/clone\_lib="2K-GTA"  
/note="Vector: ROSAFARY"

ORIGIN  
Query Match 8.8%; Score 108.6; DB 9; Length 717;  
Best Local Similarity 83.7%; Pred. No. 1.6e-18;  
Matches 123; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 55 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 114  
Db 7 CTCCTCTCAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 66  
QY 115 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 174  
Db 67 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 126  
QY 175 AGAAAAGACAATCTTTTGTGTCAG 201  
Db 127 TCCTATTCGGAAGTTCCTATTCTCTAG 153

RESULT 6  
CL256297  
LOCUS FHCRC-GT-S10-8A1 728 bp DNA linear GSS 10-FEB-2004  
DEFINITION FHCRC-GT-S10-8A1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8A1  
5', genomic survey sequence.  
ACCESSION CL256297  
VERSION CL256297.1 GI:41359945  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 728)  
Contact: Soriano P  
www.fhcr.org/labs/soriano/  
Unpublished (2003)  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at  
https://www.fhcr.org/labs/soriano/GTdb/  
Class: Gene Trap.

FEATURES source  
Location/Qualifiers  
1..728  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRC-GT-S10-8A1"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"

is available upon request at  
https://www.fhcr.org/labs/soriano/GTdb/  
Class: Gene Trap.

FEATURES source  
Location/Qualifiers  
1..717  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRC-GT-S10-8F1"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"  
/clone\_lib="2K-GTA"  
/note="Vector: ROSAFARY"

ORIGIN  
Query Match 8.8%; Score 108.6; DB 9; Length 717;  
Best Local Similarity 83.7%; Pred. No. 1.6e-18;  
Matches 123; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 55 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 114  
Db 7 CTCCTCTCAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 66  
QY 115 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 174  
Db 67 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 126  
QY 175 AGAAAAGACAATCTTTTGTGTCAG 201  
Db 127 TCCTATTCGGAAGTTCCTATTCTCTAG 153

RESULT 6  
CL256297  
LOCUS FHCRC-GT-S10-8A1 728 bp DNA linear GSS 10-FEB-2004  
DEFINITION FHCRC-GT-S10-8A1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8A1  
5', genomic survey sequence.  
ACCESSION CL256297  
VERSION CL256297.1 GI:41359945  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 728)  
Contact: Soriano P  
www.fhcr.org/labs/soriano/  
Unpublished (2003)  
Division of Basic Sciences, A2-025  
Fred Hutchinson Cancer Research Center  
1100 Fairview Ave. N., Seattle, WA 98109, USA  
Tel: 206 667 6825  
Fax: 206 667 6522  
Email: psoriano@fhcr.org  
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at  
https://www.fhcr.org/labs/soriano/GTdb/  
Class: Gene Trap.

FEATURES source  
Location/Qualifiers  
1..728  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129S4"  
/db\_xref="taxon:10090"  
/clone="FHCRC-GT-S10-8A1"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="AK7.1"

```

/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match      8.4%; Score 103.6; DB 9; Length 728;
Best Local Similarity 83.8%; Pred. No. 3.7e-17;
Matches 129; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAA 107
Db 1 GTGAGTACTCCCTCT-CAAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAA 59

QY 108 ACGAGGAGATTGATATTCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCATCCA 167
Db 60 ACGAGGAGATTGATATTCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCGCCA 119

QY 168 TCTGCTCAGAAAGACAATCTTTTGTGTCAAG 201
Db 120 AGGAAGTTCCTATTCGAGTTCCTATTCTCTAG 153

RESULT 7
CL706139      640 bp mRNA linear GSS 20-JUL-2004
LOCUS
DEFINITION
PHCRC-GT-S20-7B1 2K-GTA Mus musculus cDNA clone FHCRC-GT-S20-7B 5',
mRNA sequence.
ACCESSION
CL706139
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 640)
AUTHORS
Soriano, P.
TITLE
www.fhcr.org/labs/soriano/
JOURNAL
Unpublished (2003)
COMMENT
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and
represents the 3' insertional cDNA flanking sequence. Additional
information regarding this ES cell line and the insertion mutation
is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.
Location/Qualifiers
1..640
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S20-7B"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

FEATURES
source
Location/Qualifiers
1..579
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S8-11B"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match      7.1%; Score 88; DB 9; Length 579;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 133
Db 1 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 60

QY 134 CCCGCGGTGATGCTTTCAGGGTGGCGG 161
Db 61 CCCGCGGTGATGCTTTCAGGGTGGCGG 89

RESULT 9
CNS0091P      925 bp DNA linear GSS 03-JUN-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TEF3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.

```

```

|||||
76 TTCACCTGGCCCGCGGTGATGCTTTGAGGTGGCCCGCAACGAAGTTCCTATTCCG 135
|||||
185 ATCTTTTGTGTCAAG 201
|||||
136 AAGTTCCTATTCTCTAG 152
|||||

CG691637      579 bp DNA linear GSS 10-FEB-2004
FHCRC-GT-S8-11B1 2K-GTA Mus musculus genomic clone FHCRC-GT-S8-11B
5', genomic survey sequence.
ACCESSION
CG691637
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 579)
AUTHORS
Soriano, P.
TITLE
www.fhcr.org/labs/soriano/
JOURNAL
Unpublished (2003)
COMMENT
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and
represents the 3' insertional cDNA flanking sequence. Additional
information regarding this ES cell line and the insertion mutation
is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.
Location/Qualifiers
1..579
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S8-11B"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

FEATURES
source
Location/Qualifiers
1..579
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S8-11B"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match      7.1%; Score 88; DB 9; Length 579;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 133
Db 1 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 60

QY 134 CCCGCGGTGATGCTTTCAGGGTGGCGG 161
Db 61 CCCGCGGTGATGCTTTCAGGGTGGCGG 89

RESULT 9
CNS0091P      925 bp DNA linear GSS 03-JUN-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TEF3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.

```



```

DEFINITION  OGWIZ74TV_ZM_0.7.1.5_KB_Zea_mays_genomic_clone_ZMMBma0599M04,
              genomic survey sequence.
ACCESSION   CC721633
VERSION     CC721633.1  GI:32126409
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 645)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Contact: Cathy Whitelaw
TIGR        9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
FEATURES    Location/Qualifiers
             source
               1..645
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZMMBma0599M04"
               /clone_lib="ZM_0.7.1.5_KB"
               /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"
ORIGIN
      Query Match      4.3%; Score 53.6; DB 9; Length 645;
      Best Local Similarity 48.1%; Pred. No. 0.0021;
      Matches 152; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
      QY 218 CCGTAGAGGGCTTGACAGCAACCTGGCGATGGAGCGCAGGGTTGGTTTGTTCGGCA 277
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 419 CGCAGCGGACGCCACACACAGCGGGCGGAGCGGACGTCGGCCCTGCTGCTCTCC 360
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 278 TCGCGCGCTCTTGGCGCGATGTTTAGCTGACGATATTCGCGCGCAACGACCGCCAT 337
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 359 TCTCGTCTGCTCTGCTCGCGCTCGCAGAGCTTACGCGAGTGAGGCTGCCGCTGCACCAC 300
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 338 TCGGGAAGAGCGGTGCTGCTCGTGGCGCACAGGTGCACGCGCCACCGCGGTTGTC 397
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 299 CGCAGCAGTTCGACGGCGCGCGCTCGCGCTCTCGAGCGGGGGAGCGCGTGTGGGC 240
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 398 AGGGTGACAAAGGTCAACGCTGGTGCTACTCTCCGCTAGGCGCTCGTTGTTCCAGCAG 457
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 239 GCGGAGGAGAGTTCCGCGGGGGAGGTGGTGGCGGCGCTCGCGCGGTGGAGCGGGAG 180
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 458 AGGCGCGCGCTTCCGCGAGCAGAAATGCGGTTAGGGGTTCTAGCTGCTTCGTCGGG 517
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 179 ATGGAGGCGCGCGAGGCGGTCGCGCGGGGGTGGTGGCGGGGTGTCGCGCGGTG 120
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 518 GGGTCTGGTCCACGG 533
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 119 CGGGAGCGGTGACGG 104
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
LOCUS      BE704468
DEFINITION Sc01_01d12_R Sc01_AAFc_ECORC_cold_stressed_winter_rye_seedlings
            Secale cereale cDNA clone Sc01_01d12, mRNA sequence.
ACCESSION BE704468
VERSION   BE704468.1  GI:10092733
KEYWORDS  EST.

SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 645)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Contact: Cathy Whitelaw
TIGR        9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
FEATURES    Location/Qualifiers
             source
               1..645
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZMMBma0599M04"
               /clone_lib="ZM_0.7.1.5_KB"
               /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"
ORIGIN
      Query Match      4.3%; Score 53.6; DB 9; Length 645;
      Best Local Similarity 48.1%; Pred. No. 0.0021;
      Matches 152; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
      QY 218 CCGTAGAGGGCTTGACAGCAACCTGGCGATGGAGCGCAGGGTTGGTTTGTTCGGCA 277
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 419 CGCAGCGGACGCCACACACAGCGGGCGGAGCGGACGTCGGCCCTGCTGCTCTCC 360
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 278 TCGCGCGCTCTTGGCGCGATGTTTAGCTGACGATATTCGCGCGCAACGACCGCCAT 337
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 359 TCTCGTCTGCTCTGCTCGCGCTCGCAGAGCTTACGCGAGTGAGGCTGCCGCTGCACCAC 300
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 338 TCGGGAAGAGCGGTGCTGCTCGTGGCGCACAGGTGCACGCGCCACCGCGGTTGTC 397
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 299 CGCAGCAGTTCGACGGCGCGCGCTCGCGCTCTCGAGCGGGGGAGCGCGTGTGGGC 240
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 398 AGGGTGACAAAGGTCAACGCTGGTGCTACTCTCCGCTAGGCGCTCGTTGTTCCAGCAG 457
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 239 GCGGAGGAGAGTTCCGCGGGGGAGGTGGTGGCGGCGCTCGCGCGGTGGAGCGGGAG 180
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 458 AGGCGCGCGCTTCCGCGAGCAGAAATGCGGTTAGGGGTTCTAGCTGCTTCGTCGGG 517
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 179 ATGGAGGCGCGCGAGGCGGTCGCGCGGGGGTGGTGGCGGGGTGTCGCGCGGTG 120
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 518 GGGTCTGGTCCACGG 533
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 119 CGGGAGCGGTGACGG 104
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
LOCUS      CNS01213
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL101589
VERSION   AL101589
KEYWORDS  GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila..
            1 (bases 1 to 645)
            Genoscope.
AUTHORS

```

```

SOURCE      Secale cereale (rye)
ORGANISM    Secale cereale
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Secale.
REFERENCE   1 (bases 1 to 728)
AUTHORS    Singh,J.A., Piche,C., Couroux,P., De Moors,A., Harris,L.J.,
            Hattori,I., Ouellet,T., Robert,L.S., Spott,D. and Tinker,N.A.
            Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
            Unpublished (2000)
            Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@agr.gc.ca.
FEATURES    Location/Qualifiers
             source
               1..728
               /organism="Secale cereale"
               /mol_type="mRNA"
               /cultiivar="Puma (winter rye)"
               /db_xref="taxon:4550"
               /clone="Sc01_01d12"
               /tissue_type="leaf, crown"
               /dev_stage="seedling three-leaf stage"
               /clone_lib="Sc01_AAFc_ECORC_cold_stressed_winter_rye_seed
               ings"
               /note="Vector: Bluescript SK-/XhoI-EcoRI; Site 1: Eco RI;
               Site 2: Xho I; Sampled three-leaf seedlings treated for
               one week at 20C, 12 hrs light/day. Library made with
               Stratagene UNIZAP XR Kit/Gigapack III Gold Kit. Lambda
               library is amplified, then mass excised in SOLR cells. "
ORIGIN
      Query Match      4.3%; Score 52.8; DB 2; Length 728;
      Best Local Similarity 53.6%; Pred. No. 0.0035;
      Matches 105; Conservative 1; Mismatches 90; Indels 0; Gaps 0;
      QY 358 CTCGTGCGGACACAGGTGCACGCGCCCAACCGCGGTGTGTCAGGGTGACAAGGTCAAGCT 417
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 22 CTCTTCCTCCGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 418 GGTGCGTACTCTCTCCGCTAGGCGCTGCTGTTGTCACAGAGCGCGCGCGCTTTCGCGGA 477
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 82 GCCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 141
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 478 GCAGAAATGGCGGTAGGGGTTAGTCTGCTGCTGCTGCGGGGGTCTGCTCCACGGTAAA 537
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 142 CTGCGGTTGCGCAAGTACGTGAAGCAGCGGCTGCCCGGGGATTCGCCGCGCAGANTCAC 201
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      QY 538 GACCCCGCGGCGCAGCG 553
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 202 CGCCACGCGGCGCGCG 217
            ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
LOCUS      CNS01213
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL101589
VERSION   AL101589
KEYWORDS  GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila..
            1 (bases 1 to 645)
            Genoscope.
AUTHORS

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 gacggatcgaggatctcccc.....ctgtccctgtgtgtgtgt 100

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_phi.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_to.\*
- 11: gb\_ers.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3986	12 PCDNA3ZEO	X90639 Cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Sequence
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5082	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CVU89673	US9673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A44171	A44171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222266	AR222266 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CVU89672	US9672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1  
AR098190  
LOCUS AR098190 3853 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 5 from patent US 6074850.  
ACCESSION AR098190  
VERSION AR098190.1 GI:12807447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3853)  
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.  
TITLE Retinoblastoma fusion polypeptides  
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;  
FEATURES Location/Qualifiers  
source  
1..3853  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 3853;  
Best Local Similarity 100.0%; Pred. No. 9.4e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60  
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60  
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
RESULT 2  
AR207832  
LOCUS AR207832 3853 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 5 from patent US 6379927.  
ACCESSION AR207832  
VERSION AR207832.1 GI:21507688  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

```
Unclassified.
1 (bases 1 to 3853)
Antelman, D., Gregory, R.J. and Wills, K.N.
Retinoblastoma fusion proteins
Patent: US 6379927-A 5 30-APR-2002;
Location/Qualifiers
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 3
BD009729
LOCUS BD009729 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman, D., Gregory, R.J. and Wills, K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN, RICHARD J GREGORY, KENNETH N WILLS PC
C07H21/04, C07K5/00, A61K38/00, A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 209..862.
Location/Qualifiers
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995
LOCUS PCDNA3ZEO

REFERENCE 1 (bases 1 to 3853)
Antelman, D., Gregory, R.J. and Wills, K.N.
Retinoblastoma fusion proteins
Patent: US 6379927-A 5 30-APR-2002;
Location/Qualifiers
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

PCDNA3ZEO 4026 bp DNA linear PAT 14-FEB-2001
LOCUS PCDNA3ZEO

DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:9499972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peters, H., Hundhausen, T., Kroenke, M. and Marget, M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters, H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelistr. 5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
source
1..3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pCDNA3ZEO"
misc_feature 1..2125
/notes="cloning vector (pCDNA3) (Invitrogen)"
889..994
/notes="multiple cloning site (MCS)"
2126..2796
/notes="cloning vector (pZeoSV) (Invitrogen)"
2797..3986
/notes="cloning vector (pCDNA3)"

ORIGIN
Query Match 100.0%; Score 100; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 5
AR098191
LOCUS AR098191 4026 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman, D., Gregory, R.J. and Wills, K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
Location/Qualifiers
1..4026
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
```

```
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 6
LOCUS AR207833 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES
    source
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

RESULT 7
LOCUS BD009730 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4026
FT /organism='Unidentified'.
FEATURES
    source
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
ORIGIN

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 8
LOCUS AR098192 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES
    source
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

RESULT 9
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
    source
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

RESULT 10
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
    source
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J., and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 13-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249 /organism='Unidentified'.
FEATURES
source 1..4249
Location/Qualifiers
1..4249 /organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES
source 1..4341
Location/Qualifiers
1..4341 /organism='unidentified'
/mol_type='unassigned DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source 1..4341
Location/Qualifiers
1..4341 /organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/note='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source 1..4597
Location/Qualifiers
1..4597 /organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||

RESULT 14
AX133940
LOCUS AX133940 4840 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AX133940
VERSION AX133940.1 GI:14139881
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE Cell transfection
JOURNAL Patent: WO 0119853-A 1 22-MAR-2001;
THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES
source
1..4840
/mol_type="synthetic construct"
/db_xref="taxon:32630"
/note="This sequence is artificial and is based on well
established commercially available vectors that are cited
with their vendor within the patent application"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||

RESULT 15
BD238492
LOCUS BD238492 5053 bp DNA linear PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
of using the same.
ACCESSION BD238492
VERSION BD238492.1 GI:33048262
KEYWORDS JP 2002520000-A/18.
SOURCE synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
and Chesnut,R.W.
TITLE Expression vectors for stimulating an immune response and methods
of using the same
JOURNAL Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904,15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
FEATURES
source
Location/Qualifiers
1..5053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||

Search completed: July 14, 2005, 14:03:29
Job time : 749.127 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 gacggatcggagatctcccc.....ctgtccctgtgtgtgtgt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	1506	12 ADM41035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADM41037	Adm41037 Cytoesgal
4	100	100.0	2241	12 ADM41034	Adm41034 Human nuc
5	100	100.0	2294	12 ADM41036	Adm41036 Cytoesgal
6	100	100.0	3853	2 AAV40006	AAV40006 Plasmid p
7	100	100.0	4026	2 AAV40007	AAV40007 Plasmid p
8	100	100.0	4249	2 AAV63466	AAV63466 Plasmid p
9	100	100.0	4341	2 AAG62391	AAG62391 Vector pV
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pV
11	100	100.0	4341	6 ABN83143	Abn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AA238633	AA238633 pEP2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA seque
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rD

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 ADE21866	Ade21866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AA289476	Aa289476 Transgeni
33	100	100.0	5446	6 AAS18619	Aas18619 Renilla l
34	100	100.0	5446	6 ABL53540	Ab153540 Vector pc
35	100	100.0	5446	12 ADN36314	Adn36314 Plasmid p
36	100	100.0	5458	6 ABL58494	Ab158494 Recombina
37	100	100.0	5458	6 ABL58493	Ab158493 Recombina
38	100	100.0	5543	6 ABK88868	Abk88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ade83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	Ab158489 Recombina
42	100	100.0	5614	6 ABL58490	Ab158490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AAI66195	Aai66195 Human FSH
45	100	100.0	5651	6 ABK40237	Abk40237 DNA encod

ALIGNMENTS.

RESULT 1  
ADM41035  
ID ADM41035 standard; DNA; 1506 Bp.  
XX  
AC ADM41035;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Fungus nucleotide sequence SEQ ID NO:3.

KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; fungus; gene; ds.  
XX Unidentified.  
OS  
PN WO2004027029-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 17-SEP-2003; 2003WO-US029251.  
XX  
PR 19-SEP-2002; 2002US-0411790P.

PA (XIME-) XIMEREX INC.  
XX  
PI Beschornier WE, Sosa CE, Thompson SC;  
XX  
DR WPI; 2004-295402/27.  
XX

Engrafting foreign replacement cells within a fetal non-human mammal,  
useful in producing chimeric mammals, comprises selectively destroying  
native cells in a tissue of a fetal non-human mammal host.

Disclosure; SEQ ID NO 3; 48pp; English.

The present invention describes a method for engrafting foreign  
replacement cells within a foetal non-human mammal, which comprises  
selectively destroying native cells in a tissue of a foetal non-human  
mammal host, where the number of maternal cells of the same tissue is not  
substantially reduced, and implanting foreign replacement cells in the  
tissue of the foetal non-human mammal host, where the foreign replacement  
cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX  
SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;  
Best Local Similarity 100.0%; Pred. No. 4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 2  
ADH11349  
ID ADH11349 standard; DNA; 1600 BP.

XX  
AC ADH11349;  
DT 11-MAR-2004 (first entry)

XX  
DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX  
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
KW cell shape regulator; cell motility regulator; cell migration;  
KW cell behaviour regulator; phenotype; signal transduction pathway;  
KW signal transducing protein; signal integrator protein;  
KW neuronal regeneration; revascularisation; wound healing;  
KW chronic neurodegenerative disease; acute traumatic injury;  
KW fibrotic disease; gene; ds.

XX  
OS Unidentified.

XX  
PN WO9824810-A2.

XX  
PD 11-JUN-1998.

XX  
PF 03-DEC-1997; 97WO-EP006956.

XX  
PR 04-DEC-1996; 96GB-00025283.

XX  
PA (JANC ) JANSSEN PHARM NV.

XX  
PI Plattreuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
PI Geysen J, Bogaert THOE;

XX  
WPI; 1998-362411/31.  
DR P-PSDB; ADH11350.

XX  
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.  
PT promoting neuronal regeneration, treating chronic neuro-degenerative  
PT diseases or acute traumatic injuries.

XX  
PS Disclosure; Page 410-411; 479pp; English.

XX  
CC The present invention describes a vertebrate protein homologue of an UNC-  
CC 53 protein of Caenorhabditis elegans or a functional equivalent,  
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of  
CC cell shape, motility, or the direction of cell migration for use as a  
CC therapeutic; (7) a method for determination of whether a protein is an  
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
CC motility or the direction of migration by contacting a host cell  
CC expressing a homologue of UNC-53 and determining a change of phenotype;  
CC (8) a method for identification of vertebrate homologues of C. elegans  
CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
CC a DNA library; and (9) a method for identification of a protein which is  
CC active in the signal transduction pathway of a cell of which a vertebrate  
CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
CC antibody/homologue complex; and (iii) analysing such a complex to  
CC identify any non-antibody protein bound to the complex. UNC-53 is a  
CC signal transducing or signal integrator protein involved in controlling  
CC directionality of cell migration and cell shape in C. elegans. Vertebrate  
CC homologues of UNC-53 can be used to promote neuronal regeneration,  
CC revascularisation or wound healing, to treat chronic neurodegenerative  
CC diseases or acute traumatic injuries or fibrotic diseases. The present  
CC sequence is used in the exemplification of the present invention.

XX  
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;  
Best Local Similarity 100.0%; Pred. No. 4.1e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 3  
ADH41037  
ID ADH41037 standard; DNA; 1782 BP.

XX  
AC ADH41037;  
DT 17-JUN-2004 (first entry)

XX  
DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.

XX  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX  
OS Cytomegalovirus.

XX  
PN WO2004027029-A2.

XX  
PD 01-APR-2004.

XX  
PF 17-SEP-2003; 2003WO-US029251.

XX  
PR 19-SEP-2002; 2002US-0411790P.

XX  
PA (XIME-) XIMEREX INC.

XX  
PI Beschorner WE, Sosa CE, Thompson SC;

XX  
WPI; 2004-295402/27.

XX  
PT Engrafting foreign replacement cells within a fetal non-human mammal,  
PT useful in producing chimeric mammals, comprises selectively destroying  
PT native cells in a tissue of a fetal non-human mammal host.

XX  
PS Disclosure; SEQ ID NO 5; 48pp; English.

XX  
CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a foetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the foetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 100; DB 12; Length 1782;  
XX Best Local Similarity 100.0%; Pred. No. 4.2e-26;  
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATG 60  
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATG 60  
QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 4  
ADM41034  
ID ADM41034 standard; DNA; 2241 BP.  
XX AC ADM41034;  
XX  
XX DT 17-JUN-2004 (first entry)  
XX  
XX DE Human nucleotide sequence SEQ ID NO:2.  
XX  
XX KW engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX KW tissue transplantation; human disease study; human; Gene; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO2004027029-A2.  
XX  
XX PD 01-APR-2004.  
XX  
XX PF 17-SEP-2003; 2003WO-US029251.  
XX  
XX PR 19-SEP-2002; 2002US-0411790P.  
XX  
XX PA (XIME-) XIMEREX INC.  
XX  
XX PI Beschorner WE, Sosa CE, Thompson SC;  
XX  
XX DR WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a foetal non-human mammal host.  
XX  
XX PS Disclosure; SEQ ID NO 2; 48pp; English.  
XX  
XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the foetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 100; DB 12; Length 2241;  
XX Best Local Similarity 100.0%; Pred. No. 4.5e-26;  
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATG 60  
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATG 60  
QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 5  
ADM41036  
ID ADM41036 standard; DNA; 2294 BP.  
XX AC ADM41036;  
XX  
XX DT 17-JUN-2004 (first entry)  
XX  
XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.  
XX  
XX KW engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
XX  
XX OS Cytomegalovirus.  
XX  
XX PN WO2004027029-A2.  
XX  
XX PD 01-APR-2004.  
XX  
XX PF 17-SEP-2003; 2003WO-US029251.  
XX  
XX PR 19-SEP-2002; 2002US-0411790P.  
XX  
XX PA (XIME-) XIMEREX INC.  
XX  
XX PI Beschorner WE, Sosa CE, Thompson SC;  
XX  
XX DR WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a foetal non-human mammal host.  
XX  
XX PS Disclosure; SEQ ID NO 4; 48pp; English.  
XX  
XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the foetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for  
XX transplantation, also useful to study human diseases. The present  
XX sequence represents a nucleotide sequence given in the Sequence Listing  
XX of the present invention but not mentioned further within the  
XX specification.



```
FT CDS complement(3032..3890)
FT FT /*tag= f
FT FT /note= "AMP-ORF"
XX PN WO9821228-A1.
XX XX
XX PD 22-MAY-1998.
XX XX
XX PF 13-NOV-1997; 97WO-US021821.
XX XX
XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX XX
XX PA (CANJ-) CANJI INC.
XX PI Antelman D, Gregory RJ, Wills KN;
XX XX
XX DR WPI; 1998-297858/26.
XX XX
XX FT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX FT e.g. hyper-proliferative disease such as cancer and restenosis.
XX PS Example 1; Fig 6; 91pp; English.
XX CC This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX CC from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX CC subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX CC vector. Plasmid pCTMI has been used as a vector for the expression of
XX CC fusion proteins of the invention that comprise retinoblastoma protein
XX CC (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX CC fusion proteins, particularly expressed from gene therapy vectors, are
XX CC used to treat hyperproliferative conditions, specifically cancer
XX CC (particularly of the bladder) or restenosis. They are more effective in
XX CC repressing transcription of the E2F promoter than RB alone and cause cell
XX CC -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX CC OS field.)
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4026;
Best Local Similarity 100.0%; Pred. No. 5.3e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGAGATCTCCGATCCCGATGTCGACTTCAGTACAAATCTGCTCTGATG 60
Db |||||
Db 1 GACGATCGGAGATCTCCGATCCCGATGTCGACTTCAGTACAAATCTGCTCTGATG 60
QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db |||||
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 8
AAV63466
ID AAV63466 standard; DNA; 4249 BP.
XX AC AAV63466;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 15-FEB-1999 (first entry)
XX DE Plasmid pCTMIE.
XX KW E2F; transcription factor; human; retinoblastoma protein RB;
XX KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX KW thyroid hyperplasia; Grave's disease; psoriasis;
XX KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX KW peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX OS Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
```

Query Match 100.0%; Score 100; DB 2; Length 4249;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCGATCCCTATGTGCGACTCTCAGTACAATCTGCTCTGATG 60  
 |||||  
 Db 1 GACGGATCGGAGATCTCCGATCCCTATGTGCGACTCTCAGTACAATCTGCTCTGATG 60  
 |||||

OY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100  
 |||||  
 Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100  
 |||||

RESULT 9  
 AAO62391  
 ID AAO62391 standard; DNA; 4341 BP.  
 XX  
 AC AAO62391;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 18-NOV-1994 (first entry)  
 XX  
 DE Vector pVAC1.  
 XX  
 KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;  
 KW fusion protein; pSfi/NotI; pE1B leader; human; immunoglobulin; VH1;  
 KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;  
 KW vaccine; ss.  
 XX  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT misc\_RNA complement(1..775)  
 FT /tag= c  
 FT /note= "Claim 9"  
 FT misc\_RNA 606..780  
 FT /tag= b  
 FT /note= "Claim 8"  
 FT misc\_RNA 606..716  
 FT /tag= a  
 FT /note= "Claim 7"  
 FT  
 FT  
 XX WO9408008-A1.  
 XX  
 PD 14-APR-1994.  
 XX  
 PF 04-OCT-1993; 93WO-GB002054.  
 XX  
 PR 02-OCT-1992; 92GB-00020808.  
 XX  
 XX (MEDI-) MEDICAL RES COUNCIL.  
 XX  
 XX Hawkins RE, Russell SJ, Stevenson FK, Winter GP;  
 XX WPI; 1994-135575/16.  
 XX  
 PT Modulating immune response to a disease marker - by administering a  
 PT vector which expresses the disease marker to interact with the immune  
 PT system.  
 XX  
 PS Claim 10; Fig 7; 77pp; English.  
 XX

This sequence represents the vector pVAC1. This vector is based on the commercially available vector pRC/RSV. Leader sequences and termination signals were introduced into the vector to allow for production of fusion proteins. The vector, pSfi/NotI, was modified to replace the pE1B leader with the human immunoglobulin VH1 leader sequence that permits the encoding of an SfiI cloning site without modification of the amino acid sequence. This fragment was then cloned as an EcoRI/Blunt-HindIII fragment into NotI/Blunt-HindIII cut vector pRC/RSV to give pVAC1. The single chain Fv for an individual patient can be inserted within the VH1 leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid  
 CC vaccine and it induces a strong humoral response to the antibody moiety  
 CC in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4341;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCGATCCCTATGTGCGACTCTCAGTACAATCTGCTCTGATG 60  
 |||||  
 Db 1 GACGGATCGGAGATCTCCGATCCCTATGTGCGACTCTCAGTACAATCTGCTCTGATG 60  
 |||||

OY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100  
 |||||  
 Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100  
 |||||

RESULT 10  
 AAS17704  
 ID AAS17704 standard; DNA; 4341 BP.  
 XX  
 AC AAS17704;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Vector pVAC1 encoding a DNA vaccine.  
 XX  
 KW Cytostatic; vaccine; tetanus toxin; FrC; tumour; CTL; PCR primer; pVAC1;  
 KW ds.  
 XX  
 OS Clostridium tetani.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Cauliflower mosaic virus.  
 XX  
 PN WO200179510-A1.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 17-APR-2001; 2001WO-GB001719.  
 XX  
 PR 17-APR-2000; 2000GB-00009470.  
 XX  
 XX (CANC-) CANCER RES VENTURES LTD.  
 XX  
 XX Rice J, Stevenson F;  
 XX WPI; 2002-066370/09.  
 XX  
 XX Nucleic acid construct, useful to immunize against various diseases  
 XX including cancer, expresses the first domain of tetanus toxin FrC fused  
 XX to a disease peptide antigen to provide a vaccine.  
 XX  
 XX Disclosure; Fig 4; 71pp; English.

The invention relates to a nucleic acid construct for delivery into living cells in vivo, to induce an immune response to a disease peptide antigen, where the construct directs expression of a fusion protein comprising the peptide antigen and the first domain of FrC. Also included are a nucleic acid vector comprising the above construct, a host cell comprising the above construct or vector and a method of producing a nucleic acid construct for inducing an immune response. The method comprises identifying a nucleic acid sequence encoding a disease peptide antigen comprising epitopes characteristic of the disease, cloning the nucleic acid sequence, introducing the cloned nucleic acid into a vector which allows the antigen to be expressed as a fusion with a first domain FrC from tetanus toxin, and optionally isolating the construct from the vector. The construct or vector is used as a vaccine to induce an immune response, particularly to tumour antigens. The present sequence is vector pVAC1 which encodes a vaccine of the invention

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 11  
AEN83143  
ID AEN83143 standard; DNA; 4341 BP.  
XX  
AC AEN83143;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Plasmid pVAC1 complete sequence.  
XX  
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;  
KW cancer; B cell malignancy; ds.  
XX  
OS Synthetic.  
XX  
FN WO200240513-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-GB005142.  
XX  
PR 20-NOV-2000; 2000GB-00028319.  
XX  
PA (CANC-) 'CANCER RES VENTURES LTD.  
PI Savelyeva N, Stevenson F;  
XX  
XX WPI; 2002-500202/53.  
XX  
PT Nucleic acid construct for delivery into living cells as a vaccine,  
PT useful for treating e.g. cancer, directs the expression of a fusion  
PT protein comprising an antigen and an adjuvant sequence derived from a  
PT plant viral coat protein.  
XX  
PS Example 3; Fig 7; 84pp; English.  
XX  
CC The invention relates to a novel nucleic acid construct for inducing an  
CC immune response in vivo to an antigen, capable of directing the  
CC expression of a fusion protein that comprises an antigen and an adjuvant  
CC sequence derived from a plant viral coat protein. The construct of the  
CC invention has cytostatic and virucide activity. The nucleic acid  
CC construct is useful for inducing an immune response in a patient, for  
CC vaccinating a patient against an infectious disease caused by an antigen  
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a  
CC patient with a predisposition to cancer and for treating a patient having  
CC a B cell malignancy, where the construct is encapsidated, and optionally,  
CC a second nucleic acid sequence encoding a further immunomodulatory  
CC polypeptide is administered to the patient. The construct is also useful  
CC in medical treatment, and in the preparation of a vaccine for treating or  
CC preventing a disease state associated with the antigen. The sequence  
CC shows the complete sequence of vector pVAC1  
XX  
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 12  
AAF24901  
ID AAF24901 standard; DNA; 4597 BP.  
XX  
AC AAF24901;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.  
XX  
KW Microspheré; dihydrazide; hyaluronic acid; inflammatory response;  
KW myocardial ischemia; cardiac angiogenesis; haemophilia;  
KW vascular endothelial growth factor; VEGF; ss.  
XX  
OS Synthetic.  
XX  
FN WO200078358-A2.  
XX  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000WO-US016837.  
XX  
PR 18-JUN-1999; 99US-0140260P.  
XX  
PA (COLL-) COLLABORATIVE GROUP LTD.  
XX  
PI Chen W;  
XX  
DR WPI; 2001-071363/08.  
XX  
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial  
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic  
PT acids crosslinked to nucleic acids.  
XX  
PS Example 1; Page 36-38; 38pp; English.  
XX  
CC The specification describes a microsphere comprising dihydrazide  
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The  
CC microspheres cause reduced inflammatory responses, and have increased  
CC safety and biodegradability. The microspheres are useful for transfecting  
CC a cell of a subject and for treating a subject having myocardial  
CC ischemia, by increasing cardiac angiogenesis. They are also useful for  
CC treating haemophilia. The present sequence represents the plasmid  
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is  
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a  
CC vascular endothelial growth factor (VEGF)  
XX  
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 4; Length 4597;  
Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 13  
AAD39652

```

ID AAD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human small nuclear RNA (snRNA) DNA.
XX
KW Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN US2002058287-A1.
XX
PD 16-MAY-2002.
XX
PF 12-MAR-2001; 2001US-00804481.
XX
PR 10-MAR-2000; 2000US-0188304P.
XX
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Graaf DD, Lander ES;
XX
DR WPI; 2002-499510/53.
XX
PT New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
PS Disclosure; Fig 1; 18pp; English.
XX
CC The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
XX
AC AAF83146;
XX
DT 09-JUL-2001 (first entry)
XX
DE Complete sequence of vector pIRES/BS.
XX
KW Blastocidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
OS Synthetic.
XX
PN WO200119853-A2.

AAD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human small nuclear RNA (snRNA) DNA.
XX
KW Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN US2002058287-A1.
XX
PD 16-MAY-2002.
XX
PF 12-MAR-2001; 2001US-00804481.
XX
PR 10-MAR-2000; 2000US-0188304P.
XX
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Graaf DD, Lander ES;
XX
DR WPI; 2002-499510/53.
XX
PT New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
PS Disclosure; Fig 1; 18pp; English.
XX
CC The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
XX
AC AAF83146;
XX
DT 09-JUL-2001 (first entry)
XX
DE Complete sequence of vector pIRES/BS.
XX
KW Blastocidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
OS Synthetic.
XX
PN WO200119853-A2.
```

```

XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-GB003462.
XX
PR 11-SEP-1999; 99GB-00021418.
XX
PA (UYSH-) UNIV SHEFFIELD.
XX
PI Hollander AP, Barker MD, Kafienah W;
XX
DR WPI; 2001-290354/30.
XX
PT Novel nucleic acid molecule useful for therapeutic and cosmetic tissue
PT engineering, comprising at least a functional part of blastocidin
PT resistance gene linked through a recognition sequence, to a selected
PT gene.
XX
XX
Claim 11; Fig C; 44pp; English.
XX
CC The invention provides a nucleic acid molecule (I) comprising at least
CC the functional part of blastocidin resistance (BS) gene, or its homolog,
CC linked through a recognition sequence to at least one selected gene. (I)
CC is useful in treatment comprising: (1) providing cells/tissues transfected
CC with (I); (2) surgical administration of the cells/tissues to the patient
CC to be treated; and optionally (3) monitoring the status of the cells/
CC tissues by the patient. Therapeutic compositions comprising cells/tissues
CC transformed with (I) is useful in identifying the role of genes in
CC healthy and diseased tissue, in tissue engineering and in cosmetic
CC surgery. Tissue engineering can be used to treat arthritis, joint
CC replacement, skin grafts for burn victims, and replacement coronary
CC arteries. Cosmetic tissue surgery includes rhinoplasty. The present
CC sequence represents the nucleotide sequence of the vector pIRES/BS
CC containing the BS gene
XX
SQ Sequence 4840 BP; 1154 A; 1227 C; 1236 G; 1223 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 4; Length 4840;
Best Local Similarity 100.0%; Pred. No. 5.6e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 15
ADB33528
ID ADB33528 standard; DNA; 5015 BP.
XX
AC ADB33528;
XX
DT 04-DEC-2003 (first entry)
XX
DE Expression vector nucleotide sequence SEQ ID NO:27.
XX
KW fusion protein; amyloid precursor protein; APP; transcription factor;
KW neurotrophic; neuroprotective; APP inhibitor;
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
KW gamma-secretase; human; gene; ds.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN WO2003072041-A2.
XX
PD 04-SEP-2003.
XX
PF 23-FEB-2003; 2003WO-US005458.
```



XX 27-FEB-2002; 2002US-0360274P.  
XX (MERI ) MERCK & CO INC.  
XX  
XX Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;  
XX PI Miller MD, Register B, Shi X, Simon AU, Zuck PD;  
XX WPI; 2003-689968/65.  
XX  
XX DNA encoding a fusion protein of amyloid precursor protein, useful in  
XX screening for anti-Alzheimer agents, comprises a fused transcription  
XX factor.  
XX  
XX Disclosure; Fig 32B-E; 193pp; English.  
XX  
XX The present invention describes a DNA molecule (I) that encodes a fusion  
XX protein (FP) comprising: (i) an amino acid sequence of amyloid precursor  
XX protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a  
XX transcription factor (TF), fused in frame to the C-terminus of (i). Also  
XX described: (1) an expression vector containing (i); (2) a eukaryotic cell  
XX containing (i); and (3) methods for identifying a compound (A) that  
XX inhibits processing of APP, using the cells of (2). (i) has neurotropic and  
XX neuroprotective activities. (i) can be used to produce eukaryotic cells  
XX that express FP and are useful in screening for agents that inhibit  
XX processing of APP. The agents are potentially useful for the treatment or  
XX prevention of Alzheimer's disease. Cells that express FP can screen for  
XX inhibitors of: (a) beta- and gamma-secretases; and (b)  
XX cytoplasmic/extracellular APP signaling in a single assay. Cell-based  
XX assays may be free of interference from alpha-secretase activity and are  
XX homogeneous (no chromatography, immunoprecipitation or washing required)  
XX so well suited to high-throughput screening. The present sequence  
XX represents a plasmid nucleotide sequence from the present invention.  
XX  
SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 10; Length 5015;  
Best Local Similarity 100.0%; Pred. No. 5.6e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTCCTCTGATG 60  
|||  
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTCCTCTGATG 60  
|||  
Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100  
|||  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100  
|||

Search completed: July 14, 2005, 07:01:37  
Job time : 141.038 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_1\_100

Perfect score: 100

Sequence: 1 gacggatcggagatctccc.....ctgtccctggtgtgtgtt 100

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl1:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	60.0	602	B67169	CpG0047A Cp
2	55.6	55.6	694	BZ052929	jnr13g03.
3	55.6	55.6	696	BZ050328	jnr42c12.
4	55.6	55.6	717	BZ054067	jnr38b09.
5	53.6	53.6	348	AW409112	sali10h5 S
6	53.4	53.4	343	AL715724	AL715724 AL715724
7	53.4	53.4	345	AL714571	AL714571 AL714571
8	53.4	53.4	761	CK119397	212c09.p1
9	53.4	53.4	766	CK120360	207j04.p1
10	53.4	53.4	788	CK117844	209p08.p1
11	53.4	53.4	898	CL141237	ISB1-118J
12	53.4	53.4	899	CL140877	ISB1-118B
13	53.4	53.4	1009	CL123953	ISB1-84J1
14	53.2	53.2	814	AQ914559	nbe50049M
15	53	53.0	675	BZ051815	jnr57d03.
16	53	53.0	679	BZ052857	jnr13g03.
17	53	53.0	700	BZ050646	jnr66f08.
18	53	53.0	701	BZ052015	jnr56b03.
19	53	53.0	708	BZ054793	jnr33g03.
20	53	53.0	709	BZ053587	jnr98d01.
21	53	53.0	712	BZ054005	jnr38b09.
22	52.8	52.8	451	AQ863966	nbe50022E
23	52.6	52.6	399	AQ075099	CIT-HSP-2
24	52.4	52.4	700	BZ049113	jnr21d02.

25 52.4 52.4 708 8 BZ050047  
26 51.6 51.6 328 9 CC819886  
27 51.6 51.6 351 9 CC818492  
28 51.6 51.6 358 9 CC817661  
29 51.6 51.6 364 9 CC817805  
30 51.6 51.6 364 9 CC818511  
31 51.6 51.6 364 9 CC818574  
32 51.6 51.6 364 9 CC819049  
33 51.6 51.6 369 9 CC817069  
34 51.6 51.6 374 9 CC817074  
35 51.6 51.6 374 9 CC820036  
36 51.6 51.6 395 9 CC817652  
37 51.6 51.6 403 9 CC817682  
38 51.6 51.6 403 9 CC817837  
39 51.6 51.6 414 9 CC819240  
40 51.6 51.6 419 9 CC818384  
41 51.6 51.6 420 9 CC817834  
42 51.6 51.6 426 9 CC817720  
43 51.6 51.6 437 9 CC819820  
44 51.6 51.6 441 9 CC818421  
45 51.6 51.6 443 9 CC817769

#### ALIGNMENTS

RESULT 1  
LOCUS B67169  
DEFINITION CpG0047A CpTOWAGDNA2 Cryptosporidium parvum genomic, GSS 12-MAY-2000  
sequence.  
ACCESSION B67169  
VERSION B67169.1 GI:2642750  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum  
ORGANISM Cryptosporidium parvum  
REFERENCE 1 (bases 1 to 602)  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
MEDLINE 20183851  
PUBMED 10717299  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel.: 415 206 8846  
Fax: 415 206 3353  
Email: malaria@itsa.ucsf.edu  
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.  
Seq primer: T7  
Class: shotgun  
High quality sequence stop: 602.  
Location/Qualifiers  
1. .602  
/organism="Cryptosporidium parvum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:5807"  
/lab\_host="E. coli XL2 Blue MRF"  
/clone\_lib="CpTOWAGDNA2"  
/note="Vector: pCR-Script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)  
 . The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

## ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 41 CAGTACAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTT 100  
 Db 1 CAGTACAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTT 60

## RESULT 2

BZ052929/c  
 LOCUS jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 DEFINITION  
 ACCESSION BZ052929  
 VERSION BZ052929.1 GI:23654922  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)  
 Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr13 row: g column: 03

Seq primer: -28RPOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

## FEATURES

source

1..694  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T0100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;  
 Best Local Similarity 77.9%; Pred. No. 9e-09;  
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62

Db 324 CGGATCGATAGTCCCTGGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 265

Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

## RESULT 3

BZ050328

LOCUS jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey

DEFINITION

ACCESSION BZ050328

VERSION BZ050328.1 GI:23649718

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: c column: 12

Seq primer: -21UPPOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

FEATURES

source

1..696  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T0100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;  
 Best Local Similarity 77.9%; Pred. No. 9e-09;  
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62

Db 45 CGGATCGATAGTCCCTGGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 104

Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

## RESULT 4

BZ054067/c

LOCUS jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

DEFINITION

ACCESSION BZ054067

VERSION BZ054067.1 GI:23657216

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,



Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

## REFERENCE

1 (bases 1 to 345)  
Combra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,  
Hardelin, J. P., Weissenbach, J. and Petit, C.

## TITILE

A substracted cDNA library from the zebrafish (Danio rerio)

## JOURNAL

embryonic inner ear

## COMMENT

Unpublished (2002)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Crenieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

source

Location/Qualifiers

1..345

/organism="Danio rerio"

/mol\_type="mRNA"

/db\_xref="taxon:7955"

/clone="BN0AA007ZC02"

/tissue\_type="inner ear"

/dev\_stages="embryonic"

/clone\_lib="Danio rerio embryonic inner ear substracted

cDNA"

/note="substracted cDNA library"

## ORIGIN

Query Match 53.4%; Score 53.4; DB 1; Length 345;  
Best Local Similarity 84.5%; Pred. No. 4.8e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCGATCCCTATGTCGACTCTCAGTACATCTGCTGATGCCGCGCATAGTTAAGCCA 76  
|||||  
Db 280 TTACACCGCATATGGTGCACTCTCAGTACATCTGCTGATGCCGCGCATAGTTAAGCCA 221

QY 77 GTATCTGCTCC 87  
|||||

Db 220 GTATACACTCC 210

## RESULT 8

CK119397/c

## LOCUS

212009 pl AtM1 Arabidopsis thaliana cDNA clone MPMP2011009212

5-PRIME, mRNA sequence.

## ACCESSION

CK119397

CK119397.1 GI:47829713

EST.

## SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 761)

Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

Contact: Birgit Kersten

Plant Protein Chip Group, Department Leirach

Max-Planck-Institute for Molecular Genetics

Imenstr. 73 D-14195 Berlin, Germany

Tel: +49(0)30/84131648

Fax: +49(0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 761 Std Error: 0.00

Plate: 212 row: O column: 9

Seq primer: PQE65.

Location/Qualifiers

1..761

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Columbia"

/db\_xref="GABI:954234"

/db\_xref="taxon:3702"

/clone="MPMP2011009212"  
/tissue\_type="inflorescence meristem"  
/dev\_stage="about one week after bolting"  
/lab\_host="E. coli SCS-1/pSE111"  
/clone\_lib="AtM1"  
/note="Vector: PQE-30NAST-attB (AY386205); Site 1: SalI;  
Site 2: NotI; About 1 week after bolting, cDNA synthesis  
using SuperscriptTM-system (Invitrogen) with an  
oligo(dT)-primer containing NotI restriction site and a  
SalI adapter. The main library (plate numbers begin with  
1) of 38,000 clones was rearrayed into the sublibrary  
(plate numbers begin with 201) containing 5,000 putative  
expression clones. Average insert size is 1 kb. Note: The  
rearrayed sublibrary (plate numbers begin with 201) was  
sequenced. Library generation and sequencing was granted  
in context of GABI-LAPP; data are also accessible at  
https://gabi.rzpd.de"

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 761;  
Best Local Similarity 84.5%; Pred. No. 5.6e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76  
|||||  
Db 674 TTACACCGCATATGGTGCACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 615

QY 77 GTATCTGCTCC 87  
|||||

Db 614 GTATACACTCC 604

## RESULT 9

CK120360/c

## LOCUS

207104 pl AtM1 Arabidopsis thaliana cDNA clone MPMP2011J04207

5-PRIME, mRNA sequence.

## ACCESSION

CK120360

CK120360.1 GI:47830676

EST.

## SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 766)

Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

Contact: Birgit Kersten

Plant Protein Chip Group, Department Leirach

Max-Planck-Institute for Molecular Genetics

Imenstr. 73 D-14195 Berlin, Germany

Tel: +49(0)30/84131648

Fax: +49(0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 766 Std Error: 0.00

Plate: 207 row: J column: 4

Seq primer: PQE65.

Location/Qualifiers

1..766

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Columbia"

/db\_xref="GABI:953059"

/db\_xref="taxon:3702"

/clone="MPMP2011J04207"

/tissue\_type="inflorescence meristem"

/dev\_stage="about one week after bolting"

/lab\_host="E. coli SCS-1/pSE111"

/clone\_lib="AtM1"

/note="Vector: PQE-30NAST-attB (AY386205); Site 1: SalI;

Site\_2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearayed into the sublibrary (plate numbers begin with 201) containing 5,000 putative expression clones. Average insert size is 1 kb. Note: The rearayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 766;  
Best Local Similarity 84.5%; Pred. No. 5.6e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTGATGCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 76  
Db 679 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 620

Qy 77 GTATCTGCTCC 87  
Db 619 GTATACACTCC 609

## RESULT 10

CK117844/c  
LOCUS CK117844 788 bp mRNA linear EST 01-JUN-2004  
DEFINITION 209p08.p1 Atm1 Arabidopsis thaliana cDNA clone MPMPGP2011P08209  
5-PRIME, mRNA sequence.

ACCESSION CK117844  
VERSION CK117844.1 GI:47828160  
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 788)  
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.  
TITLE Generation of a cDNA expression library from Arabidopsis inflorescence meristem  
JOURNAL Unpublished (2003)  
COMMENT Contact: Birgit Kersten  
Plant Protein Chip Group, Department Lehrach  
Max-Planck-Institute for Molecular Genetics  
Innestr. 73, D-14195 Berlin, Germany  
Tel: +49(0)30/84131648  
Fax: +49(0)30/84131128  
Email: Kersten@molgen.mpg.de

Insert Length: 788 Std Error: 0.00  
Plate: 209 row: P column: 8  
Seq primer: pQ65.

## FEATURES

Location/Qualifiers  
1..788  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="GABI:953578"  
/db\_xref="taxon:3702"  
/clone="MPMPGP2011P08209"  
/tissue\_type="inflorescence meristem"  
/dev\_stage="about one week after bolting"  
/lab\_host="E. coli SCS-1/pSE111"  
/clone\_lib="AtM1"

/note="Vector: pQ8-30NAST-attB (AY386205); Site 1: SalI; Site 2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearayed into the sublibrary (plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The rearayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 788;  
Best Local Similarity 84.5%; Pred. No. 5.7e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTGATGCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 76  
Db 514 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 455

Qy 77 GTATCTGCTCC 87  
Db 454 GTATACACTCC 444

## RESULT 11

CL141237/c  
LOCUS ISB1-118J17\_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118J17,  
DEFINITION genomic survey sequence.  
ACCESSION CL141237  
VERSION CL141237.1 GI:40634872  
KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 898)  
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,

Mardis,E. and Wilson,R.  
TITLE A physical map of the xenopus tropicalis genome  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K Wilson

Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert Length: 75000 Std Error: 0.00  
Seq primer: T7 TAATACGACTCACTATAGG

Class: BAC ends  
High quality sequence start: 4  
High quality sequence stop: 742.

## FEATURES

Location/Qualifiers  
1..898  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8364"  
/clone="ISB1-118J17"  
/clone\_lib="ISB1"  
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"

## ORIGIN

Query Match 53.4%; Score 53.4; DB 9; Length 898;  
Best Local Similarity 84.5%; Pred. No. 5.8e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTGATGCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 76  
Db 195 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATCGCGCATAGTTAAGCCA 136

Qy 77 GTATCTGCTCC 87  
Db 135 GTATACACTCC 125

## RESULT 12

CL140877/c

```

LOCUS       CL140877               899 bp    DNA        linear    GSS 05-JAN-2004
DEFINITION   ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
genomic survey sequence.
ACCESSION    CL140877
VERSION      CL140877.1   GI:40634512
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 899)
AUTHORS      Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE        A physical map of the xenopus tropicalis genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 75000 Std Error: 0.00
              Seq primer: T7 TAATACGACTCTACTATAGGG
              Class: BAC ends
              High quality sequence start: 4
              High quality sequence stop: 681.
              Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"

ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 899;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTTAAGCCA 76
DB 195 TTCACACCGCATATGGTGCATCTCTCAGTACAACTGCTGTGATGCCGATAGTTAAGCCA 136

QY 77 GTATCTGCTCC 87
DB 135 GTATACACTCC 125

RESULT 13
LOCUS       CL123953/c
DEFINITION   ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
genomic survey sequence.
ACCESSION    CL123953
VERSION      CL123953.1   GI:40617588
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 1009)
AUTHORS      Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE        A physical map of the xenopus tropicalis genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 75000 Std Error: 0.00
              Seq primer: T7 TAATACGACTCTACTATAGGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
Location/Qualifiers
  1..1009
  /organism="Xenopus tropicalis"
  /mol_type="genomic DNA"
  /db_xref="taxon:8364"
  /clone="ISB1-84J15"
  /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
  Library Segment 1"

ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTTAAGCCA 76
DB 252 TTCACACCGCATATGGTGCATCTCTCAGTACAACTGCTGTGATGCCGATAGTTAAGCCA 193

QY 77 GTATCTGCTCC 87
DB 192 GTATACACTCC 182

RESULT 14
LOCUS       AQ914559               814 bp    DNA        linear    GSS 02-DEC-1999
DEFINITION   nbeb0049M21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
cultivar-group) genomic clone nbeb0049M21r, genomic survey
sequence.
ACCESSION    AQ914559
VERSION      AQ914559.1   GI:6511075
KEYWORDS     GSS.
SOURCE       Oryza sativa (japonica cultivar-group)
              Oryza sativa (japonica cultivar-group)
              Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzaceae; Oryza.
              1 (bases 1 to 814)
              Wing,R.A. and Dean,R.A.
              A BAC End Sequencing Framework to Sequence the Rice Genome
              Unpublished (1998)
              Contact: Wing RA
              Clemson University Genomics Institute
              Clemson University
              100 Jordan Hall, Clemson, SC 29634, USA
              Tel: 864 656 7288
              Fax: 864 656 4293
              Email: rwing@clemson.edu
              Seq primer: GGAACAGCTATGACCATG
              Class: BAC ends
              High quality sequence start: 21
              High quality sequence stop: 361.
              Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nbeb0049M21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /note="vector: pBACindigo; Site:1: EcorI; Site:2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```



Barle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, the Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center ([www.genome.clemson.edu](http://www.genome.clemson.edu))."

## ORIGIN

Query Match 53.2%; Score 53.2; DB 8; Length 814;  
 Best Local Similarity 78.0%; Pred. No. 6.7e-08;  
 Matches 64; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 7 TCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCCGAT 66  
 |||||  
 DB 279 TGGGCGGATTTTCACACCGCATATGTCGACTCTCAGTACAATCTGCTCTGATGCCGAT 338  
 |||||

QY 67 AGTTAAGCCAGTATCTGCTCCC 88  
 |||||

DB 339 AGTTAAGCCAGCCCGCACCC 360  
 |||||

## RESULT 15

BZ051815  
 LOCUS  
 DEFINITION jnr57d03.b1 B.oleracea001 Brassica oleracea genomic, genomic survey sequence.  
 BZ051815 675 bp DNA linear GSS 09-OCT-2002  
 VERSION BZ051815.1 GI:23652690  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 675)  
 Delehaunty, K., Fewell, G., Fulton, L., McComb, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
 Whole genome shotgun reads from Brassica oleracea  
 Unpublished (2002)  
 Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 Plate: jnr57 row: d column: 03  
 Seq primer: -21UPpOT forward  
 Class: shotgun  
 High quality sequence start: 29  
 High quality sequence stop: 94.

## FEATURES

source  
 1..675  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;  
 Best Local Similarity 75.6%; Pred. No. 7.6e-08;

Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CGGATCGGGAGATCTCCCGATCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62  
 |||||  
 DB 53 CGGNACGATAGGTCCCTGGACTAGTTATGTTGCGACTCTCAGTACAATCTGCTCTGATGCC 112  
 |||||

QY 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||

DB 113 GCATAGTTAAGCCAGCCCGCACCC 138  
 |||||

Search completed: July 14, 2005, 23:22:51  
 Job time : 952.146 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_4141\_4241

Perfect score: 101  
Sequence: 1 acagcaagggggaggattgg.....ccagctggggctcaggggg 101

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	3986	12	PCDNA3ZEO
2	101	100.0	4597	6	X90639 Cloning vec
3	101	100.0	5082	6	AX060344 Sequence
4	101	100.0	5082	6	A91754 Sequence 10
5	101	100.0	5432	6	BD085110 Vertebrat
6	101	100.0	5432	6	BD234590 Screening
7	101	100.0	5432	6	AX026821 Sequence
8	101	100.0	5446	6	BD195386 Compositi
9	101	100.0	5446	6	AX319694 Sequence
10	101	100.0	5590	12	AB038602 Cloning v
11	101	100.0	5639	12	AX437643 Expressio
12	101	100.0	5651	6	AX211282 Sequence
13	101	100.0	5651	6	AX349366 Sequence
14	101	100.0	5731	6	AX202478 Sequence
15	101	100.0	5995	6	AX685746 Sequence
16	101	100.0	6084	12	CGA575208
17	101	100.0	6109	12	TRU90717
18	101	100.0	6148	6	BD181637
19	101	100.0	6149	6	AX342685 Sequence
					BD181638 Novel mel

20	101	100.0	6149	6	AX342686
21	101	100.0	6180	6	AX207724
22	101	100.0	6186	6	AX211281
23	101	100.0	6186	6	AX349365
24	101	100.0	6195	6	BD168975
25	101	100.0	6213	6	AX211283
26	101	100.0	6213	6	AX349369
27	101	100.0	6238	6	BD168966
28	101	100.0	6253	6	AR031374
29	101	100.0	6253	6	BD009742
30	101	100.0	6277	12	AV437644
31	101	100.0	6331	12	EVPCWVPA1
32	101	100.0	6333	12	EVPCWVPA3
33	101	100.0	6334	6	AX665478
34	101	100.0	6335	12	EVPCWVPA2
35	101	100.0	6338	6	BD134374
36	101	100.0	6338	6	AR428934
37	101	100.0	6340	6	AX207733
38	101	100.0	6365	6	AX513181
39	101	100.0	6394	12	AF416990
40	101	100.0	6404	6	BD267665
41	101	100.0	6411	6	AX207725
42	101	100.0	6411	6	AX207729
43	101	100.0	6420	6	BD267666
44	101	100.0	6436	6	AX207740
45	101	100.0	6439	6	AR240214

ALIGNMENTS

RESULT 1  
PCDNA3ZEO  
LOCUS PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995  
DEFINITION Cloning vector pcdna3zEO DNA.  
ACCESSION X90639  
VERSION X90639.1 GI:949972  
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Peters,H., Hundhausen,T., Kroenke,M. and Marget,M.  
TITLE A new small sized high-level eukaryotic expression vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3986)  
AUTHORS Peters,H.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,  
Michaelistr.5, D- 24105 Kiel, FRG  
COMMENT Related sequences: M21295 and K03104.  
FEATURES  
source  
1..3986  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
/plasmid="pcDNA3ZEO"  
1..2125  
/note="cloning vector (pcDNA3) (Invitrogen)"  
889..994  
/note="multiple cloning site (MCS)"  
2126..2796  
/note="cloning vector (PZeoSV) (Invitrogen)"  
2797..3986  
/note="cloning vector (pcDNA3)"

ORIGIN

Query Match 100.0%; Score 101; DB 12; Length 3986;  
Best Local Similarity 100.0%; Pred.No.2.5e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ACAGCAAGGGGGAGGATTGGGAGACACATAGCAGGCATGCTGGGATCGCGTGGGCTCTA 60  
|||||

Db 1204 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 1263

Qy 61 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 101  
 |||||  
 Db 1264 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 1304  
 |||||

RESULT 2  
 AX060344  
 LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001  
 DEFINITION Sequence 3 from Patent WO0078358.  
 ACCESSION AX060344  
 VERSION AX060344.1 GI:12405832  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.

REFERENCE 1  
 AUTHORS Chen, W.  
 TITLE Hyaluronic acid microspheres for sustained gene transfer  
 JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;  
 The Collaborative Group, Ltd. (US)

FEATURES  
 source  
 1. 4597  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN  
 Query Match 100.0%; Score 101; DB 6; Length 4597;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 60  
 |||||  
 Db 1780 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 1839  
 |||||

Qy 61 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 101  
 |||||  
 Db 1840 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 1880  
 |||||

RESULT 3  
 A91754  
 LOCUS A91754 5082 bp DNA circular PAT 22-JAN-2000  
 DEFINITION Sequence 10 from Patent WO9824810.  
 ACCESSION A91754  
 VERSION A91754.1 GI:6740671  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unclassified.

REFERENCE 1 (bases 1 to 5082)  
 AUTHORS Bogaert, T.A. and Deraeymaeker, M.  
 TITLE VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS  
 JOURNAL Patent: WO 9824810-A 10 11-JUN-1998;  
 BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEMYAEKER MARC (BE)

FEATURES  
 source  
 1. 5082  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 100.0%; Score 101; DB 6; Length 5082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 60  
 |||||  
 Db 2874 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 2933  
 |||||

Qy 61 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 101  
 |||||  
 Db 2934 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 2974  
 |||||

RESULT 4  
 BD085110  
 LOCUS BD085110 5082 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Vertebrate homologues of UNC-53 protein of C elegans.  
 ACCESSION BD085110  
 VERSION BD085110.1 GI:22630720  
 KEYWORDS JP 2001522222-A/8.  
 SOURCE unidentified  
 ORGANISM unclassified.

REFERENCE 1 (bases 1 to 5082)  
 AUTHORS Platteuw, C.J., Arjol, C.M.B., Deraeymaeker, M., Verhasselt, P.,  
 Pujol, N.J.R., Luc, Maertens, J.S., Luyten, W., Geerts, H.,  
 Vandekerckhove, J.S., Geysen, J. and Bogaert, T.A.O.E.  
 TITLE Vertebrate homologues of UNC-53 protein of C elegans  
 JOURNAL Patent: JP 2001522222-A 8 13-NOV-2001;  
 JANSSEN PHARMACEUTICA NV

COMMENT  
 OS Unidentified  
 PN JP 2001522222-A/8  
 PD 13-NOV-2001  
 PF 03-DEC-1997 JP 1998525231  
 PR 04-DEC-1996 GB 9625283.8  
 PI CHRIST JULES PLATTEUW, CARLOS MANUEL BUESA ARJOL, MARC PI  
 DERAEMYAEKER,  
 PI PETER VERHASSELT, NATHALIE JEANNE RAYMONDE PUJOL, LUC PI  
 JACQUES SIMON MAERTENS,  
 PI WALTER LUYTEN, HUGO GEERTS, JOEL STEFAAN VANDEKERCKHOVE, JOHAN  
 PI GEYSEN,  
 PI THIERRY ANDRE OLIVIER EDDY BOGAERT  
 PC C12N15/12, C12N5/10, C12N15/85, C07K14/435, C07K16/18, A61K38/17,  
 PC A61K49/00,  
 PC C12Q1/02, G01N33/53  
 CC Strandedness: Double;  
 CC Topology: Circular;  
 CC Vertebrate homologues of UNC-53 protein of C elegans FH Key  
 Location/Qualifiers  
 FT source  
 1. 5082  
 /organism="Unidentified".  
 Location/Qualifiers  
 1. 5082  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 100.0%; Score 101; DB 6; Length 5082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 60  
 |||||  
 Db 2874 ACAGCAAGGGGAGGATTGGGAAGACATACAGGCATGCTGGGATCGGTGGGCTCTA 2933  
 |||||

Qy 61 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 101  
 |||||  
 Db 2934 TGGCTTCTAGCGGGAAGAACACAGCTGGGCTCTAGGGG 2974  
 |||||

RESULT 5  
 BD234590  
 LOCUS BD234590 5432 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Screening assay of Abeta-peptide.  
 ACCESSION BD234590  
 VERSION BD234590.1 GI:33044360  
 KEYWORDS JP 2002531141-A/2.  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.

```
REFERENCE 1 (bases 1 to 5432)
AUTHORS Peraus,G.
TITLE Screening assay of Abeta-peptide
JOURNAL Patent: JP 2002531141-A 2 24-SEP-2002;
COMMENT AVENTIS PHARMA DEUTSCHLAND GMBH
OS Artificial Sequence
PN JP 2002531141-A/2
PD 24-SEP-2002
PF 27-NOV-1999 JP 2000586944
PR 07-DEC-1998 DE 198 56 261.6
PI GISELA PERAUS
PC C12N15/09,A01K67/033,A61K45/00,A61P25/28,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12Q1/37,C12Q1/68,C12N15/00,C12N5/00 CC Description
of Artificial Sequence: Mutagen
FH Key Location/Qualifiers
FT source 1..5432
FT 1..5432 /organism='Artificial Sequence'.
FEATURES
source
1..5432
/organism="synthetic construct"
/mol type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
DB 1190 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 1249
QY 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
DB 1250 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 1290
RESULT 6
AX026821
LOCUS Peraus,G. 5432 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 9 from Patent DE19856261.
ACCESSION AX026821
VERSION AX026821.1 GI:10187947
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peraus,G.
JOURNAL Patent: DE 19856261-C 9 30-MAR-2000;
HOBCHST MARION ROUSSEL DE GMBH (DE)
FEATURES
source
1..5432
/organism="synthetic construct"
/mol type="unassigned DNA"
/db_xref="taxon:32630"
/note="Mutagen"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
DB 1190 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 1249
QY 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
DB 1250 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 1290
```

```
RESULT 7
BD195386 5446 bp DNA linear PAT 17-JUL-2003
LOCUS Composition and methods for administering Pneumococcal DNA.
DEFINITION BD195386
ACCESSION BD195386
VERSION BD195386.1 GI:33005156
KEYWORDS JP 2002514061-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 5446)
AUTHORS Briles,D.E., Mcdaniel,L.S. and Curiel,D.T.
TITLE Composition and methods for administering Pneumococcal DNA
JOURNAL Patent: JP 2002514061-A 3 14-MAY-2002;
UNIVERSITY OF ALABAMA AT BIRMINGHAM
COMMENT OS Unidentified
PN JP 2002514061-A/3
PD 14-MAY-2002
PF 04-DEC-1997 JP 1998525895
PR 04-DEC-1996 US 08/759505
PI DAVID E BRILES, LARRY S MCDANIEL, DAVID T CURIEL PC
C12P21/06,C12N15/00,C07H21/02,C07H21/04
CC Strandedness: Single;
CC Topology: Linear;
CC Composition and methods for administering Pneumococcal DNA FH
Key Location/Qualifiers
FT source 1..5446
FT 1..5446 /organism='Unidentified'.
FEATURES
source
1..5446
/organism="unidentified"
/mol type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACACCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
DB 1204 ACACCAAGGGGAGGATTGGGAAGACATAGCAGCATGCTGGGATCGGTGGGCTCTA 1263
QY 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
DB 1264 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 1304
RESULT 8
AX319694 5446 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 5 from Patent WO0181614.
DEFINITION AX319694
ACCESSION AX319694
VERSION AX319694.1 GI:17901350
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Leng,J.
TITLE Cell proliferation assay
JOURNAL Patent: WO 0181614-A 5 01-NOV-2001;
Chemicon International (US)
FEATURES
source
1..5446
/organism="synthetic construct"
/mol type="unassigned DNA"
/db_xref="taxon:32630"
/note="pcDNA3 vector sequence"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
```

```
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGAAGACAATACAGGCATGCTGGGATGCGGTGGGCTCTA 60
    |||
Db 1204 ACAGCAAGGGGAGGATTGGAAGACAATACAGGCATGCTGGGATGCGGTGGGCTCTA 1263
    |||

Qy 61 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 101
    |||
Db 1264 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 1304
    |||

RESULT 9
AB038602/c
LOCUS AB038602 5590 bp DNA circular SYN 24-JAN-2004
DEFINITION Cloning vector pLCPVRGNR104 DNA, complete sequence.
ACCESSION AB038602
VERSION AB038602.1 GI:13094141
KEYWORDS Cloning vector pLCPVRGNR104
SOURCE Cloning vector pLCPVRGNR104
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 Hashinaka,K.
AUTHORS Synthetic Autonomous Vectors Based on Palindromic Sequences of
TITLE Parvovirus B19
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 5590)
AUTHORS Hashinaka,K.
TITLE Direct Submision
JOURNAL Submitted (21-FEB-2000) Kazuya Hashinaka, Miyazaki Medical College,
Department of Biochemistry, 5200 Kihara, Kiyotake, Miyazaki
889-1692, Japan (E-mail:hashinaka@post1.miyazaki-med.ac.jp,
Tel:81-985-85-0985, Fax:81-985-85-2401)
FEATURES
source
1. 5590
/organism="Cloning vector pLCPVRGNR104"
/mol_type="other DNA"
/db_xref="taxon:117920"
/focus
12. 485
/organism="B19 virus"
/mol_type="other DNA"
/db_xref="taxon:10798"
/notes="synonym:Parvovirus B19"
3369. 3846
/organism="B19 virus"
/mol_type="other DNA"
/db_xref="taxon:10798"
/notes="synonym:Parvovirus B19"
15. 397
complement(1115..2635)
/gene="rsGFP-neor"
complement(1115..2635)
/gene="rsGFP-neor"
/codon_start=1
/transl_table=11
/product="red-shift green fluorescent protein-fused
neomycin phosphotransferase"
/protein_id="BAB32740.1"
/db_xref="GI:13094142"
/translacion="MASKGELFTGVVPILVELDGDVNGHKFVSQGEQDATYKGLT
LKFICTTKGVPVMTLTTCYGVQCFSRYPDLGMKRDFFKSAPEGYVOERTIFPK
DDGNYKTRHNIEGDSVLVNRIELKIDFKEDGNILGHLEYNYSNNVYIMADKQK
GIKYVFTAGITHGMBELVYNGAIEODGLHAGSPAARVERLFYDWAQTIICSDA
MVLLEFVTAAGITHGMBELVYNGAIEODGLHAGSPAARVERLFYDWAQTIICSDA
AVRLSAGRPVLVFKTDLSCALNELOEARLSWLATGVPVCAALDVVTEAGRWL
LLGVPVQDILLSHLAPAEKVISIMADAMRLRLTDPATCFPDHQAHRIRERPRMEA
GLVDDDDDEEHQGDIALATRDIAELGGEWADRFLVLYGIAAPDSQRIAFYRLLDFF"
3455. 3837
complement(4834..5493)
/gene="Cmr"
complement(4834..5493)

repeat_region
gene
CDS
repeat_region
gene
CDS
```

```
/gene="Cmr"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="BAB32741.1"
/db_xref="GI:13094143"
/translacion="MEKKTGYTTVDISQWHEKHEFFAFOSVACQTNQTVHLDITAF
LKTWKKKKHFFYFAFIFILARLNAHTEFRMAKDGELVWDSVHPCTYVFEQETEF
SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGF TENMFFVFSANPWVSTFDLNV
ANMNFAPVFTMGKTYTQGDVKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDWQGG
A"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5590;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGAAGACAATACAGGCATGCTGGGATGCGGTGGGCTCTA 60
    |||
Db 925 ACAGCAAGGGGAGGATTGGAAGACAATACAGGCATGCTGGGATGCGGTGGGCTCTA 866
    |||

Qy 61 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 101
    |||
Db 865 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 825
    |||

RESULT 10
AY437643
LOCUS AY437643 5639 bp DNA circular SYN 10-NOV-2003
DEFINITION Expression vector pcGlobin 2, complete sequence.
ACCESSION AY437643
VERSION AY437643.1 GI:38155839
KEYWORDS Expression vector pcGlobin 2
SOURCE Expression vector pcGlobin 2
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 5639)
AUTHORS Ro,H., Kim,E.J. and Rhee,M.
TITLE A new vector system, pcGlobin 2 for in vitro synthesized RNA
injection into zebrafish embryos
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5639)
AUTHORS Ro,H., Kim,E.J. and Rhee,M.
TITLE Direct Submision
JOURNAL Submitted (14-OCT-2003) Department of Biology, College of Natural
Sciences, Chungnam National University, 305-764, Daejeon 305-764,
Korea
FEATURES
source
1. 5639
/organism="Expression vector pcGlobin 2"
/mol_type="other DNA"
/db_xref="taxon:254096"
/notes="eukaryotic expression vector for zebrafish embryo
microinjection; derivative of pcDNA3"
complement(4643..5503)
/codon_start=1
/product="beta-lactamase"
/protein_id="AAR12689.1"
/db_xref="GI:38155840"
/translacion="MSIQHFRVALIPFFAAFCPLVPFAHPETLVKVKDAEDOLGARVGY
IEDLSGKILSFRRPEERPFMMSTFKVLLCGAVLSRIDAGQEQLGRIHYSONDLVE
YSPVTRKHLTDGMTVRELCSAITSNDTAAALLLTITGGIKELTAFLHNGMDVTRL
DRWPELNAEIPNDRDITTPMVAMATIRKLITGLLETLTASRQQLIDWMEADKVRGL
LRGALPAGWPIADKSGAGERSGIITAAALGPDGKPSRIVIVITGTSQATMDERNQIA
EIGASLIKWH"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5639;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGAAGACAATACAGGCATGCTGGGATGCGGTGGGCTCTA 60
    |||
```

Db 1397 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1456

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
|||||  
Db 1457 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1497  
|||||

RESULT 11  
LOCUS AX2111282 5651 bp DNA linear PAT 06-SEP-2001  
DEFINITION Sequence 6 from Patent WO0158493.  
ACCESSION AX2111282  
VERSION AX2111282.1 GI:155233691  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Schambye,H.T., Andersen,K.V., van den Hazel,B., Christiansen,J. and Jeppesen,C.B.  
TITLE Conjugates of follicle stimulating hormones  
JOURNAL Patent: WO 0158493-A 6 16-AUG-2001;  
Maxygen Aps (DK)  
FEATURES  
source Location/Qualifiers  
1..5651  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
1231..1617  
/note="Coding sequence for human FSH-beta"

exon

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5651;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 1826 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1885  
|||||

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
|||||  
Db 1886 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1926  
|||||

RESULT 12  
LOCUS AX349366 5651 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 4 from Patent WO0202597.  
ACCESSION AX349366  
VERSION AX349366.1 GI:18615329  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Okkels,J.S., Jensen,A.D. and van den Hazel,B.C.  
TITLE Peptide extended glycosylated polypeptides  
JOURNAL Patent: WO 0202597-A 4 10-JAN-2002;  
Maxygen Aps (DK); Maxygen Holdings Ltd (KY)  
FEATURES  
source Location/Qualifiers  
1..5651  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
1231..1617  
/note="Coding sequence for human FSH-beta"

exon

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5651;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 1826 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1885  
|||||

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
|||||  
Db 1886 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1926  
|||||

RESULT 13  
LOCUS AX202478 5731 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 66 from Patent WO0152620.  
ACCESSION AX202478  
VERSION AX202478.1 GI:15392206  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.  
TITLE Methods and compositions to modulate expression in plants  
JOURNAL Patent: WO 0152620-A 66 26-JUL-2001;  
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)  
FEATURES  
source Location/Qualifiers  
1..5731  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="2C7-SID"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5731;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 1906 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1965  
|||||

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
|||||  
Db 1966 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 2006  
|||||

RESULT 14  
LOCUS AX685746 5995 bp DNA linear PAT 29-MAR-2003  
DEFINITION Sequence 5 from Patent WO02102854.  
ACCESSION AX685746  
VERSION AX685746.1 GI:29371751  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Thomassen-Wolf,E., Borges,E., Yayon,A. and Rom,E.  
TITLE Antibodies that block receptor protein tyrosine Kinase activation, Methods of screening for and uses thereof  
JOURNAL Patent: WO 02102854-A 5 27-DEC-2002;  
MorphoSys AG (DE); ProChon Biotech Ltd. (IL)  
FEATURES  
source Location/Qualifiers  
1..5995  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
963..1670  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/db\_xref="GI:29371752"

CDS

/translation="DPBEPKSCDKTHCPCPAPBELLGGSVFLFPKPKDXTLMISRT  
PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPRBEQYNSTYRVVSVLTVLHQDWL  
NGKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYITLPFSDELTRKNQVSLTCLVKG  
YPSDIAVWESNGQPENNYKTTPPVLDSGDSFFLYSKLTVDKSRWQQGNVPSQSVWHE  
ALHNHYTQKSLSLSPGK"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5995;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 ACAGCAAGGGGAGATTGGGAAGCATATACAGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 1870 ACAGCAAGGGGAGATTGGGAAGCATATACAGCATGCTGGGGATGCGGTGGGCTCTA 1929  
|||||  
  
Qy 61 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 101  
|||||  
Db 1930 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 1970  
|||||

RESULT 15  
GGA575208  
LOCUS GGA575208 6084 bp DNA circular SYN 03-JUL-2003  
DEFINITION Expression vector pCLGFPa.  
ACCESSION AJ575208  
VERSION AJ575208.1 GI:32451228  
KEYWORDS  
SOURCE  
ORGANISM Expression vector pCLGFPa  
other sequences; artificial sequences; vectors.  
REFERENCE 1  
AUTHORS Scaal,M., Gros,J., Lesbros,C. and Marcelle,C.  
TITLE In ovo electroporation of avian somites  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 6084)  
AUTHORS Marcelle,C.  
TITLE Direct Submission  
JOURNAL Submitted (13-JUN-2003) Marcelle C., Lgpd, Institut de Biologie du  
Developpement, Campus de Luminy Case 907 F-13288 Marseille, 13288,  
FRANCE

FEATURES

Location/Qualifiers  
1..6084  
/organism="Expression vector pCLGFPa"  
/mol\_type="other DNA"  
/db\_xref="taxon:236984"  
/focus  
12..397  
/organism="unidentified cytomegalovirus"  
/mol\_type="other DNA"  
/db\_xref="taxon:205912"  
398..1652  
/organism="Gallus gallus"  
/mol\_type="other DNA"  
/db\_xref="taxon:9031"  
1653..1733  
/organism="Oryctolagus cuniculus"  
/mol\_type="other DNA"  
/db\_xref="taxon:9986"  
1782..1996  
/organism="Bos taurus"  
/mol\_type="other DNA"  
/db\_xref="taxon:9913"  
2442..2836  
/organism="Simian virus 40"  
/mol\_type="other DNA"  
/db\_xref="taxon:10633"  
2908..3723  
/organism="Aequorea victoria"  
/mol\_type="other DNA"  
/db\_xref="taxon:6100"  
12..399  
740..887  
/gene="beta actin"

promoter 740..887  
/gene="beta actin"  
1734..1776  
/notes="multiple cloning site; MCS  
under control of CMV enhancer/chicken beta-actin promoter"  
  
polyA\_signal 1782..1996  
/notes="bovine growth hormone"  
enhancer 2442..2836  
misc\_feature 2908..3723  
/notes="eGFP  
under control of SV40 promoter"  
polyA\_signal 3757..3887  
/notes="late"  
  
ORIGIN  
  
Query Match 100.0%; Score 101; DB 12; Length 6084;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 ACAGCAAGGGGAGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 1951 ACAGCAAGGGGAGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 2010  
|||||  
  
Qy 61 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 101  
|||||  
Db 2011 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 2051  
|||||

Search completed: July 14, 2005, 14:03:30  
Job time : 757.618 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_4141\_4241  
Perfect score: 101  
Sequence: 1 acagcaagggggaggattgg.....ccagctggggctctaggggg 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	3482	2	ADH11353
2	101	100.0	4597	4	Aaf24901 Nucleotid
3	101	100.0	4825	13	ADRI12380
4	101	100.0	5015	10	ADB33528 Expressio
5	101	100.0	5082	2	ADH11417
6	101	100.0	5218	12	ADM97811
7	101	100.0	5302	12	ADI34681
8	101	100.0	5425	2	ADH11233
9	101	100.0	5431	6	ABN86685
10	101	100.0	5431	10	ADE21866
11	101	100.0	5431	12	ADO5277
12	101	100.0	5432	3	AAZ89476
13	101	100.0	5446	2	AAV38297
14	101	100.0	5446	6	AA518619
15	101	100.0	5446	6	ABL53540
16	101	100.0	5446	12	ADN36314
17	101	100.0	5458	6	ABL58494
18	101	100.0	5458	6	ABL58493
19	101	100.0	5614	6	ABL58489
20	101	100.0	5614	6	ABL58490

21	101	100.0	5651	5	AAI66195
22	101	100.0	5651	6	ABK40237
23	101	100.0	5695	6	ABL58492
24	101	100.0	5695	6	ABL58491
25	101	100.0	5695	8	ABT40262
26	101	100.0	5695	8	ADA89054
27	101	100.0	5695	10	ADG74306
28	101	100.0	5731	4	AAD11615
29	101	100.0	5821	12	ADM97787
30	101	100.0	5864	6	AAI44423
31	101	100.0	5864	6	AAI44424
32	101	100.0	6082	8	AAI56212
33	101	100.0	6082	8	AAI56211
34	101	100.0	6082	8	AAI56210
35	101	100.0	6085	8	AAI56213
36	101	100.0	6094	8	AAI56215
37	101	100.0	6097	8	AAI56214
38	101	100.0	6100	6	ABK96469
39	101	100.0	6135	6	ABK96470
40	101	100.0	6148	6	ABK15579
41	101	100.0	6149	6	ABK15580
42	101	100.0	6180	4	AAI13062
43	101	100.0	6186	5	AAI66194
44	101	100.0	6186	6	ABK40236
45	101	100.0	6195	6	ABK51585

ALIGNMENTS

RESULT 1  
ADH11353  
ID ADH11353 standard; DNA; 3482 BP.  
XX  
AC ADH11353;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Vertebrate UNC-53 protein homologue related nucleotide sequence.  
XX  
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans; cell shape regulator; cell motility regulator; cell migration;  
KW cell behaviour regulator; phenotype; signal transduction pathway;  
KW signal transducing protein; signal integrator protein;  
KW neuronal regeneration; revascularisation; wound healing;  
KW chronic neurodegenerative disease; acute traumatic injury;  
KW fibrotic disease; gene; ds.  
XX  
OS Unidentified.  
XX  
PN WO9824810-A2.  
XX  
PD 11-JUN-1998.  
XX  
XX 03-DEC-1997; 97WO-EP006956.  
XX  
PR 04-DEC-1996; 96GB-00025283.  
XX  
PA (JANC ) JANSSEN PHARM NV.  
XX  
PI Platteauw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
PI Geyssen J, Bogaert TAOE,  
XX  
XX WPI; 1998-362411/31.  
DR P-PSDB; ADH11354.  
XX  
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g. promoting neuronal regeneration, treating chronic neuro-degenerative diseases or acute traumatic injuries.  
XX  
PS Disclosure; Page 414-417; 479pp; English.  
XX

The present invention describes a vertebrate protein homologue of an UNC-53 protein of *Caenorhabditis elegans* or a functional equivalent, derivative or precursor of UNC-53. Also described: (1) a cDNA sequence encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a nucleic acid which hybridises to the cDNA of (1); (3) vector comprising the cDNA as in (1); (4) a host cell containing the vector as in (3); (5) a transgenic cell, tissue or animal comprising the vector as in (3); (6) a compound identified as an enhancer or inhibitor of the regulation of a cell shape, motility, or the direction of cell migration for use as a therapeutic; (7) a method for determination of whether a protein is an inhibitor or enhancer of regulation of cell behaviour, growth, shape or motility or the direction of migration by contacting a host cell expressing a homologue of UNC-53 and determining a change of phenotype; (8) a method for identification of vertebrate homologues of *C. elegans* unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to a DNA library; and (9) a method for identification of a protein which is active in the signal transduction pathway of a cell of which a vertebrate homologue of UNC-53 is a component comprising: (i) contacting an extract of a cell with an antibody to the UNC-53 homologue; (ii) identifying an antibody/homologue complex; and (iii) analysing such a complex to identify any non-antibody protein bound to the complex. UNC-53 is a signal transducing or signal integrator protein involved in controlling directionality of cell migration and cell shape in *C. elegans*. Vertebrate homologues of UNC-53 can be used to promote neuronal regeneration, revascularisation or wound healing, to treat chronic neurodegenerative diseases or acute traumatic injuries or fibrotic diseases. The present sequence is used in the exemplification of the present invention.

	Query Match	100.0%	Score 101;	DB 2;	Length 3482;
	Best Local Similarity	100.0%;	Pred. No. 1.6e-22;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACAGCAGGGGGAGGATTGGGAACACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	60		
Db	1274	ACAGCAGGGGGAGGATTGGGAACACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	1333		
Qy	61	TGGCTTCTAGGCGGGAAGAACACAGCTGGGGCTCTAGGGG	101		
Db	1334	TGGCTTCTAGGCGGGAAGAACACAGCTGGGGCTCTAGGGG	1374		

RESULT 2	
AAF24901	
ID	AAF24901 standard; DNA; 4597 BP.
XX	
XX	AAF24901;
AC	
XX	
DT	20-APR-2001 (first entry)
XX	
XX	Nucleotide sequence of the plasmid pCDNA3.1/GS.
DE	
XX	
KW	Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW	myocardial ischemia; cardiac angiogenesis; haemophilia;
KW	vascular endothelial growth factor; VEGF; ss.
XX	
XX	Synthetic.
OS	

PT	Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT	ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT	acids crosslinked to nucleic acids.
XX	
XX	Example 1; Page 36-38; 38pp; English.
PS	
XX	The specification describes a microsphere comprising dihydrazide
CC	derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The
CC	microspheres cause reduced inflammatory responses, and have increased
CC	safety and biodegradability. The microspheres are useful for transfecting
CC	a cell of a subject and for treating a subject having myocardial
CC	ischemia, by increasing cardiac angiogenesis. They are also useful for
CC	treating haemophilia. The present sequence represents the plasmid
CC	pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC	crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC	vascular endothelial growth factor (VEGF)
XX	
XX	Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;
XX	
XX	Query Match 100.0%; Score 101; DB 4; Length 4597;
XX	Best Local Similarity 100.0%; Pred. No. 1.7e-22;
XX	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	
QY	1 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGCTCTA 60
DB	1780 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGCTCTA 1839
XX	
QY	61 TGGCTTCTGAGGGCGAAGACACAGCTGGGGCTCTAGGGGG 101
DB	1840 TGGCTTCTGAGGGCGAAGACACAGCTGGGGCTCTAGGGGG 1880
XX	
XX	RESULT 3
AD	ADRL2380
ID	ADRL2380 standard; DNA; 4825 BP.
XX	
AC	ADRL2380;
XX	
DT	21-OCT-2004 (first entry)
XX	
DE	Vector pMCP1.
XX	
KW	ss; cytostatic; VEGF modulator; angiogenesis inhibitor;
KW	UTR-dependent expression; vascular endothelial growth factor;
KW	untranslated region; cancer; angiogenesis; vector.
XX	
OS	Synthetic.
XX	
XX	WO2004065561-A2.
XX	
PD	05-AUG-2004.
XX	
PF	21-JAN-2004; 2004WO-US001643.
XX	
PR	21-JAN-2003; 2003US-0441637P.
XX	
PA	(PTCT-) PTC THERAPEUTICS INC.
XX	
PI	Cao L, Trifillis P;
XX	
DR	WPI; 2004-571681/55.
XX	
PT	Identifying modulators of untranslated region-dependent expression of a
PT	VEGF gene, useful for treating cancer, comprises contacting a compound
PT	with a cell or translation mixture containing a reporter gene linked to a
PT	VEGF gene UTR.
XX	
PS	Example; SEQ ID NO 94; 251pp; English.
XX	
CC	A method of identifying (M1) a compound that modulates untranslated
CC	region-dependent expression of a vascular endothelial growth factor
CC	(VEGF) gene comprises contacting a member of a library of compounds with
CC	a cell or cell-free translation mixture containing a reporter gene

	Query Match	100.0%;	Score 101;	DB 4;	Length 4597;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-22;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACAGCAAGGGGAGGATTTGGAAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	60		
Db	1780	ACAGCAAGGGGAGGATTTGGAAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	1839		
Qy	61	TGGCTTCTAGGCGGAAAGAACACAGCTGGGGCTCTAGGGGG	101		
Db	1840	TGGCTTCTAGGCGGAAAGAACACAGCTGGGGCTCTAGGGGG	1880		

RESULT 3	
ADR12380	
ID	ADR12380 standard; DNA; 4825 BP.
XX	
AC	ADR12380;
XX	
DT	21-OCT-2004 (first entry)
XX	
DE	Vector pMCPI.
XX	
KW	ss; cytostatic; VEGF modulator;
KW	UTR-dependent expression; vascular
KW	untranslated region; cancer; angi
XX	
OS	Synthetic.
XX	
PN	WO2004065561-A2.
XX	
PD	05-AUG-2004.
XX	
PF	21-JAN-2004; 2004WO-US001643.
XX	
PR	21-JAN-2003; 2003US-0441637P.
XX	
PA	(PTCT-) PTC THERAPEUTICS INC.
XX	
PI	Cao L, Trifillis P;
XX	
DR	WPI: 2004-571681/55.

CC operably linked to an untranslated region (UTR) of the VEGF gene, and  
 CC detecting expression of the reporter gene. A compound is identified as  
 CC modulator if the level of expression of the reporter gene in the presence  
 CC of the compound is altered as compared to that in the absence of the  
 CC compound or in the presence of a control. Compounds identified by M1 are  
 CC useful for treating, preventing or ameliorating cancer or its symptoms,  
 CC and/or for inhibiting, angiogenesis. This sequence corresponds to the  
 CC vector pMCP1, a mammalian expression vector designed to integrate into  
 CC the genome at sites containing the PRT recombination site using the flp  
 CC recombinase.

XX SQ Sequence 4825 BP; 1236 A; 1135 C; 1204 G; 1250 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 13; Length 4825;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACCAAGGGGAGGATTGGGAACACATACAGGCATCTGGGGATCGGTGGGCTCTA 60  
 |||||  
 Db 2601 ACACCAAGGGGAGGATTGGGAACACATACAGGCATCTGGGGATCGGTGGGCTCTA 2660  
 QY 61 TGGCTTCTGAGCGGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
 |||||  
 Db 2661 TGGCTTCTGAGCGGGAAGAACACAGCTGGGGCTCTAGGGGG 2701

RESULT 4  
 ADB33528  
 ID ADB33528 standard; DNA; 5015 BP.

XX AC ADB33528;

XX DT 04-DEC-2003 (first entry)

XX DE Expression vector nucleotide sequence SEQ ID NO:27.

XX KW fusion protein; amyloid precursor protein; APP; transcription factor;  
 KW neurotropic; neuroprotective; APP inhibitor;  
 KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;  
 KW gamma-secretase; human; gene; ds.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO2003072041-A2.

XX PD 04-SEP-2003.

XX PF 23-FEB-2003; 2003WO-US005458.

XX PR 27-FEB-2002; 2002US-0360274P.

XX PA (MERI ) MERCK & CO INC.

XX PI Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;  
 PI Miller MD, Register B, Shi X, Simon AJ, Zuck PD;

XX DR WPI; 2003-689968/65.

XX PT DNA encoding a fusion protein of amyloid precursor protein, useful in  
 PT screening for anti-Alzheimer agents, comprises a fused transcription  
 PT factor.

XX PS Disclosure; Fig 32B-F; 193pp; English.

XX CC The present invention describes a DNA molecule (I) that encodes a fusion  
 CC protein (PP) comprising: (i) an amino acid sequence of amyloid precursor  
 CC protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a  
 CC transcription factor (TF), fused in frame to the C-terminus of (i). Also  
 CC described: (1) an expression vector containing (I); (2) a eukaryotic cell  
 CC containing (I); and (3) methods for identifying a compound (A) that  
 CC inhibits processing of APP, using the cells of (2). (I) has neurotropic and  
 CC neuroprotective activities. (I) can be used to produce eukaryotic cells

CC that express PP and are useful in screening for agents that inhibit  
 CC processing of APP. The agents are potentially useful for the treatment or  
 CC prevention of Alzheimer's disease. Cells that express PP can screen for  
 CC inhibitors of: (a) beta- and gamma-secretases; and (b)  
 CC cytoplasmic/extracellular APP signaling in a single assay. Cell-based  
 CC assays may be free of interference from alpha-secretase activity and are  
 CC homogeneous (no chromatography, immunoprecipitation or washing required)  
 CC so well suited to high-throughput screening. The present sequence  
 CC represents a plasmid nucleotide sequence from the present invention.

XX SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 10; Length 5015;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACCAAGGGGAGGATTGGGAACACATACAGGCATCTGGGGATCGGTGGGCTCTA 60  
 |||||  
 Db 1190 ACACCAAGGGGAGGATTGGGAACACATACAGGCATCTGGGGATCGGTGGGCTCTA 1249  
 QY 61 TGGCTTCTGAGCGGGAAGAACACAGCTGGGGCTCTAGGGGG 101  
 |||||  
 Db 1250 TGGCTTCTGAGCGGGAAGAACACAGCTGGGGCTCTAGGGGG 1290

RESULT 5  
 ADH11417  
 ID ADH11417 standard; DNA; 5082 BP.

XX AC ADH11417;

XX DT 11-MAR-2004 (first entry)

XX DE Plasmid pCB201 nucleotide sequence.

XX KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
 KW cell shape regulator; cell motility regulator; cell migration;  
 KW cell behaviour regulator; phenotype; signal transduction pathway;  
 KW signal transducing protein; signal integrator protein;  
 KW neuronal regeneration; revascularisation; wound healing;  
 KW chronic neurodegenerative disease; acute traumatic injury;  
 KW fibrotic disease; human; gene; ds.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 XX CDS 1028..2149  
 XX /\*tag= a

XX PN WO9824810-A2.

XX PD 11-JUN-1998.

XX PF 03-DEC-1997; 97WO-EP006956.

XX PR 04-DEC-1996; 96GB-00025283.

XX PA (JANC ) JANSSEN PHARM NV.

XX PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
 PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

XX PL Geysen J, Bogaert TAOB;

XX DR WPI; 1998-362411/31.

XX DR P-PSDB; ADH11424.

XX PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.  
 PT promoting neuronal regeneration, treating chronic neuro-degenerative  
 PT diseases or acute traumatic injuries.

XX PS Claim 96; SEQ ID NO 10; 479pp; English.

CC The present invention describes a vertebrate protein homologue of an UNC-  
CC 53 protein of *Caenorhabditis elegans* or a functional equivalent,  
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
CC encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a  
CC nucleic acid which hybridizes to the cDNA of (1); (3) vector comprising  
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)  
CC a compound identified as an enhancer or inhibitor of the regulation of  
CC cell shape, motility, or the direction of cell migration for use as a  
CC therapeutic; (7) a method for determination of whether a protein is an  
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
CC motility or the direction of migration by contacting a host cell  
CC expressing a homologue of UNC-53 and determining a change of phenotype;  
CC (8) a method for identification of vertebrate homologues of *C. elegans*  
CC UNC-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
CC a DNA library; and (9) a method for identification of a protein which is  
CC active in the signal transduction pathway of a cell of which a vertebrate  
CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
CC antibody/homologue complex; and (iii) analysing such a complex to  
CC identify any non-antibody protein bound to the complex. UNC-53 is a  
CC signal transducing or signal integrator protein involved in controlling  
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate  
CC homologues of UNC-53 can be used to promote neuronal regeneration,  
CC revascularisation or wound healing, to treat chronic neurodegenerative  
CC diseases or acute traumatic injuries or fibrotic diseases. The present  
CC sequence is used in the exemplification of the present invention.

XX  
SQ Sequence 5082 BP; 1164 A; 1365 C; 1311 G; 1242 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5082;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60  
Db 2874 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 2933

Qy 61 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG 101  
Db 2934 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG 2974

RESULT 6  
ADM97811  
ID ADM97811 standard; DNA; 5218 BP.  
XX  
AC ADM97811;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE & X UAS beta-lactamase vector SEQ ID NO: 64.  
XX  
KW enzyme; sensor cell; signal transduction detection system; promoter;  
KW targeting sequence; targeted drug; ds; vector.  
XX  
OS Synthetic.  
XX Unidentified.  
XX WO2004031415-A2.  
XX  
PN 15-APR-2004.  
XX  
PD 05-SEP-2003; 2003WO-US028078.  
XX  
PF 05-SEP-2002; 2002US-0408297P.  
XX  
PR (VERT-) VERTEX PHARM INC.  
XX  
PA Whitney MA, Zeh K, Sanders PS;  
XX  
PI WPI; 2004-330208/30.  
XX  
DR  
XX

PT Developing a sensor cell, useful in determining the activity of a target  
PT gene and in developing therapeutic drugs, comprises providing cells  
PT comprising a signal transduction detection system and introducing DNA  
PT construct into cells.  
XX  
PS Example 7; Page 231-234; 234pp; English.  
XX  
CC The present invention relates to a method of developing a sensor cell,  
CC for determining the activity of a target gene in the cell, which  
CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing  
CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
CC vector used in the exemplification of the invention.

XX  
SQ Sequence 5218 BP; 1231 A; 1361 C; 1335 G; 1291 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5218;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60  
Db 4177 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 4236

Qy 61 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG 101  
Db 4237 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG 4277

RESULT 7  
ADI34681  
ID ADI34681 standard; DNA; 5302 BP.  
XX  
AC ADI34681;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Nucleotide sequence of plasmid pcDNA6/Biotag (TM) -D-TOPO.  
XX  
KW Recombinational cloning; recombination; topoisomerase;  
KW fusion protein production; ds.  
XX  
OS Synthetic.  
XX WO2004005482-A2.  
XX  
PN 15-JAN-2004.  
XX  
PD 08-JUL-2003; 2003WO-US021339.  
XX  
PF 08-JUL-2002; 2002US-0393756P.  
XX  
PR 19-JUL-2002; 2002US-0396627P.  
XX  
PR 10-OCT-2002; 2002US-0417172P.  
XX  
PA (INVI-) INVITROGEN CORP.  
XX  
PI Bennett RP;  
XX  
DR WPI; 2004-091356/09.  
XX  
PT New isolated nucleic acid molecules having one or more recombination  
PT sites and encoding an amino acid sequence tag, useful for recombinational  
PT and/or topoisomerase-mediated cloning methods for producing fusion  
PT proteins.  
XX  
PS Example 1; Fig 11A-B; 135pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC one or more recombination sites, and one or more nucleic acid sequences

CC which encode an amino acid sequence tag. (1) can also comprise one or  
CC more topoisomerase recognition sites and/or one or more topoisomerases.  
CC The amino acid sequence tag is an amino acid sequence that is capable of  
CC being post-translationally modified, and is an amino acid sequence that  
CC is capable of being post-translationally modified by biotinylation,  
CC attachment of 4-phosphopantetheine, attachment of lipoid acid or  
CC attachment of flavins, and is an amino acid sequence that is capable of  
CC being biotinylated, wherein the amino acid sequence that is capable of  
CC being biotinylated is all or a portion of the Klebsiella pneumoniae  
CC oxalacetate decarboxylase a subunit, all or a portion of the  
CC Propionibacterium shermanii transcarboxylase 1.3S subunit, or all or a  
CC portion of the Escherichia coli biotin carboxyl carrier protein component  
CC of acetyl-CoA carboxylase. The methods and compositions of the present  
CC invention are useful for identifying, concentrating, purifying and/or  
CC producing fusion proteins that comprise an amino acid sequence tag. The  
CC nucleic acid molecules can also be used in recombinational cloning and/or  
CC topoisomerase-mediated cloning methods to produce polynucleotide  
CC constructs which encode the fusion proteins. The present sequence  
CC represents the nucleotide sequence of a plasmid pcDNA6/Biotag(TM)-D-TOPO  
SQ Sequence 5302 BP; 1254 A; 1361 C; 1349 G; 1338 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5302;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGGTGGGCTCTTA 60  
Db 1425 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGGTGGGCTCTTA 1484  
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101  
Db 1485 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1525

RESULT 8  
ADH11233  
ID ADH11233 standard; DNA; 5425 BP.  
AC ADH11233;  
DT 11-MAR-2004 (first entry)  
DE Vertebrate UNC-53 protein homologue related nucleotide sequence.  
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
KW cell shape regulator; cell motility regulator; cell migration;  
KW cell behaviour regulator; phenotype; signal transduction pathway;  
KW signal transducing protein; signal integrator protein;  
KW neuronal regeneration; revascularisation; wound healing;  
KW chronic neurodegenerative disease; acute traumatic injury;  
KW fibrotic disease; gene; ds.  
OS Unidentified.  
XX  
XX WO9824810-A2.  
XX  
XX 11-JUN-1998.  
XX  
XX 03-DEC-1997; 97WO-EP006956.  
XX  
XX 04-DEC-1996; 96GB-00025283.  
XX  
XX (JANC ) JANSSEN PHARM NV.  
XX  
XX Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
PI Puijol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
PI Geyseu J, Bogaert TAOE;  
XX  
XX WPI; 1998-362411/31.  
DR P-PSDB; ADH11234.  
XX  
XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

PT promoting neuronal regeneration, treating chronic neuro-degenerative  
PT diseases or acute traumatic injuries.  
XX  
XX Disclosure; Page 231-237; 479pp; English.  
CC The present invention describes a vertebrate protein homologue of an UNC-  
CC 53 protein of Caenorhabditis elegans or a functional equivalent,  
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)  
CC a compound identified as an enhancer or inhibitor of the regulation of  
CC cell shape, motility, or the direction of cell migration for use as a  
CC therapeutic; (7) a method for determination of whether a protein is an  
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
CC motility or the direction of migration by contacting a host cell  
CC expressing a homologue of UNC-53 and determining a change of phenotype;  
CC (8) a method for identification of vertebrate homologues of C. elegans  
CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
CC a DNA library; and (9) a method for identification of a protein which is  
CC active in the signal transduction pathway of a cell of which a vertebrate  
CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
CC antibody/homologue complex; and (iii) analysing such a complex to  
CC identify any non-antibody protein bound to the complex. UNC-53 is a  
CC signal transducing or signal integrator protein involved in controlling  
CC directionalities of cell migration and cell shape in C. elegans. Vertebrate  
CC homologues of UNC-53 can be used to promote neuronal regeneration,  
CC revascularisation or wound healing, to treat chronic neurodegenerative  
CC diseases or acute traumatic injuries or fibrotic diseases. The present  
CC sequence is used in the exemplification of the present invention.  
SQ Sequence 5425 BP; 1250 A; 1463 C; 1420 G; 1292 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5425;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGGTGGGCTCTTA 60  
Db 3217 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGGTGGGCTCTTA 3276  
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101  
Db 3277 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 3317  
RESULT 9  
ABN86685  
ID ABN86685 standard; DNA; 5431 BP.  
XX  
XX AC ABN86685;  
XX  
XX 05-NOV-2002 (first entry)  
XX  
XX Nucleotide sequence of a pcDNA3 plasmid vector.  
XX  
XX Major histocompatibility complex; MHC; antigen presenting cell; APC;  
KW antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;  
KW cancer; viral infection; ds.  
XX  
XX Synthetic.  
XX  
XX WO200261113-A2.  
XX  
XX 08-AUG-2002.  
XX  
XX 01-FEB-2002; 2002WO-US002598.  
XX  
XX 01-FEB-2001; 2001US-0265334P.  
XX  
XX (UVJO ) UNIV JOHNS HOPKINS.

XX Wu T, Hung C;  
XX WPI; 2002-619261/66.  
XX  
XX Nucleic acid molecule encoding a fusion polypeptide that promotes  
XX processing via the Major Histocompatibility Complex class I pathway  
XX and/or promotes activity of an antigen presenting cell, useful as vaccine  
XX for cancer and viral infections.  
XX  
XX Claim 24; Page 22-23; 127pp; English.  
XX  
XX The invention relates to a new nucleic acid molecule (I) encoding a  
XX fusion polypeptide useful as a vaccine composition. (I) comprises a first  
XX nucleic acid sequence encoding a first polypeptide or peptide that  
XX promotes processing via the Major Histocompatibility Complex (MHC) class  
XX I pathway (MHC-I-pp) and/or promotes development or activity of an  
XX antigen presenting cell (APC). The second nucleic acid sequence is linked  
XX in frame to the first nucleic acid sequence or to a linker nucleic acid  
XX sequence and encodes an antigenic polypeptide or peptide. The methods and  
XX compositions of the present invention are useful as therapeutic vaccine  
XX for cancer and for major viral infections, such as hepatoma and cervical  
XX cancer, that cause morbidity and mortality. They can also be used in  
XX treating animal diseases, such as equine herpesvirus, bovine viruses,  
XX Marek's disease, retroviral and lentiviral diseases and rabies, in the  
XX veterinary medicine context. The present sequence represents the  
XX nucleotide sequence of a pcDNA3 plasmid vector  
XX  
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 6; Length 5431;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 60  
Db 1189 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 1248  
Qy 61 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGCTCTAGGGGG 101  
Db 1249 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGCTCTAGGGGG 1289  
RESULT 10  
ADE21866  
ID ADE21866 standard; DNA; 5431 BP.  
XX  
XX ADE21866;  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Plasmid vector pcDNA3 nucleotide sequence SEQ ID NO:8.  
XX  
XX chimeric fusion; translocation; antigenic; cytostatic; immunotherapy;  
XX gene therapy; cancer; tumour; gene; ds.  
XX  
XX Synthetic.  
XX  
XX WO2003085085-A2.  
XX  
XX 16-OCT-2003.  
XX  
XX 04-APR-2003; 2003WO-US010235.  
XX  
XX 04-APR-2002; 2002US-00115440.  
XX  
XX (UJO ) UNIV JOHNS HOPKINS.  
XX  
XX Wu T, Hung C;  
XX  
XX WPI; 2003-877027/81.  
XX  
XX New nucleic acid encoding a chimeric fusion or fusion polypeptide

PT comprising a first domain with a translocation polypeptide, and a second  
PT domain with an antigen having at least one antigenic peptide, useful for  
PT treating cancer.  
XX  
XX Disclosure; SEQ ID NO 8; 68pp; English.  
XX  
XX The present invention describes a nucleic acid (I) encoding a chimeric  
XX fusion or fusion polypeptide comprising a first domain with a  
XX translocation polypeptide, and a second domain comprising an antigen  
XX having at least one antigenic peptide. Also described: (1) an expression  
XX vector comprising (I) operatively linked to a promoter and optionally, to  
XX one or more regulatory elements that enhance expression of the nucleic  
XX acid in a cell; (2) a particle comprising (I) or the expression vector;  
XX (3) a cell that has been modified to comprise (I) or the expression  
XX vector; (4) a chimeric polypeptide comprising a first domain with a  
XX translocation polypeptide, and a second domain comprising an antigen  
XX having at least one antigenic peptide; (5) a pharmaceutical composition  
XX capable of inducing or enhancing an antigen specific immune response,  
XX comprising (I), expression vector, particle, cell, cell of the particle,  
XX or the chimeric polypeptide; and a carrier or excipient; (6) inducing or  
XX enhancing an antigen specific immune response by administering the  
XX composition described above; (7) increasing the number of CD8 + CTLs  
XX specific for a selected desired antigen in a subject by administering the  
XX composition described above; and (8) inhibiting the growth of a tumour in  
XX a subject by administering the composition described above. (I) has  
XX cytostatic activity, and can be used in immunotherapy, and gene therapy.  
XX The nucleic acids (I), compositions and methods are useful for treating  
XX cancer. The present sequence represents a plasmid vector nucleotide  
XX sequence which is used in the exemplification of the present invention.  
XX  
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 10; Length 5431;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 60  
Db 1189 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 1248  
Qy 61 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGCTCTAGGGGG 101  
Db 1249 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGCTCTAGGGGG 1289  
RESULT 11  
ADO05277  
ID ADO05277 standard; DNA; 5431 BP.  
XX  
XX ADO05277;  
XX  
XX 29-JUL-2004 (first entry)  
XX  
XX pcDNA3 plasmid vector.  
XX  
XX Translocation domain; bacterial toxin; exotoxin A domain II; ETA;  
XX major histocompatibility complex; MHC class I; vaccine; immune response;  
XX CD8+ cytotoxic T lymphocyte; CTL; tumour; E7 antigen; pcDNA3 plasmid; ds.  
XX  
XX Synthetic.  
XX  
XX US2004086845-A1.  
XX  
XX 06-MAY-2004.  
XX  
XX 04-APR-2002; 2002US-00115440.  
XX  
XX 20-OCT-1999; 99US-00421608.  
XX 09-FEB-2000; 2000US-00501097.  
XX 20-OCT-2000; 2000WO-US041422.  
XX 04-APR-2001; 2001US-0281003P.  
XX  
XX (WUTT/) WU T.

PA (HUNG/) HUNG C.  
XX  
PI Wu T, Hung C;  
XX  
DR WPI; 2004-356187/33.  
XX  
XX Novel chimeric polypeptide e.g., Pseudomonas aeruginosa exotoxin A domain  
PT I1/human papilloma virus-16 E7 peptide useful for inducing or enhancing  
PT antigen specific immune response, or for inhibiting growth of tumor in  
PT subject.  
XX  
XX Disclosure; SEQ ID NO 8; 48pp; English.  
XX  
XX The invention relates to nucleic acid encoding a chimeric polypeptide  
CC comprising a translocation domain of a bacterial toxin and at least one  
CC antigenic peptide. The preferred translocation domain is domain II of  
CC Pseudomonas aeruginosa exotoxin A (ETA(dII)) and the preferred antigen is  
CC human papilloma virus type 16 (HPV-16) E7 which is a model tumour  
CC antigen. The antigenic peptide comprises an epitope that binds to and is  
CC presented on the cell surface by major histocompatibility complex (MHC)  
CC class I proteins. The nucleic acid of the invention is useful as vaccine  
CC composition for enhancing antigen specific immune response, increasing  
CC the number of CD8+ cytotoxic T lymphocytes (CTLs) and for inhibiting the  
CC growth of a tumour. The present sequence is pCDNA3 plasmid vector used in  
CC the invention.  
XX  
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 12; Length 5431;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
DB 1189 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1248  
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101  
DB 1249 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1289  
RESULT 12  
AAZ89476  
ID AAZ89476 standard; DNA; 5432 BP.  
XX  
AC AAZ89476;  
XX  
DT 22-JUN-2000 (first entry)  
XX  
DE Transgenic APP DNA #2.  
XX  
KW APP; amyloid precursor protein; gamma-secretase; neuroprotective;  
KW nontropic; transgenic; Alzheimer's disease; Down's syndrome; ds.  
XX  
OS Synthetic.  
XX  
PN DE19856261-C1.  
XX  
PD 30-MAR-2000.  
XX  
PF 07-DEC-1998; 98DE-01056261.  
XX  
PR 07-DEC-1998; 98DE-01056261.  
XX  
PA (HMR) ) HOECHST MARION ROUSSEL DEUT GMBH.  
XX  
PI Peraus G;  
XX  
XX WPI; 2000-258119/23.  
XX  
PT Detection of gamma-secretase by detection of A-beta peptide useful for  
PT determining gamma-secretase activity and for identifying inhibitors.  
XX

PS Claim 30; Page 7-8; 16pp; German.  
XX  
XX This invention describes a novel method for the detection of human gamma-  
CC secretase by detection of a partial protein formed by cleavage of a  
CC fusion protein encoded by a transgene containing a first nucleotide  
CC sequence which encodes a protein comprising the amino acid sequence (A)  
CC and a second nucleotide sequence which encodes a signal peptide. The  
CC products of the invention have neuroprotective and nontropic activity.  
CC The method is used to detect activity of gamma-secretase. The transgene  
CC and/or vectors are useful for the production of a transgenic cell or C.  
CC elegans. Transgenic C. elegans is useful in a method for the  
CC determination of gamma-secretase activity. The transgenic C. elegans is  
CC also useful in a method to identify inhibitors of the gamma-secretase  
CC activity. The methods and transgenes are useful in research of  
CC Alzheimer's disease. Inhibitors of gamma-secretase are useful in  
CC control/treatment of Alzheimer's and possibly Down's syndrome. This  
CC sequence encodes a transgenic amyloid precursor protein (APP) which is  
CC described in the method of the invention  
XX  
SQ Sequence 5432 BP; 1251 A; 1410 C; 1390 G; 1381 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 3; Length 5432;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
DB 1190 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 1249  
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101  
DB 1250 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1290  
RESULT 13  
AAV38297  
ID AAV38297 standard; DNA; 5446 BP.  
XX  
AC AAV38297;  
XX  
DT 17-OCT-2003 (revised)  
DT 26-OCT-1998 (first entry)  
XX  
DE Plasmid pCDNA3.  
XX  
KW Plasmid pCDNA3; pneumococcal surface protein A; PspA; infection;  
KW Streptococcus pneumoniae; sepsis; otitis media; meningitis; bacteraemia;  
KW pneumonia; vaccine; genetic immunisation; ss.  
XX  
OS Human herpesvirus 5.  
OS Chimeric.  
XX  
PN WO9824927-A1.  
XX  
PD 11-JUN-1998.  
XX  
PF 04-DEC-1997; 97WO-US022847.  
XX  
PR 04-DEC-1996; 96US-00759505.  
XX  
PA (UYAL-) UNIV ALABAMA.  
XX  
PI Briles DE, McDaniel LS, Curriel DT;  
XX  
DR WPI; 1998-333343/29.  
XX  
PT Plasmid containing pneumococcal epitope for expression in eukaryotic  
PT cells - useful for eliciting immunological response to pneumococcal  
PT infection or sepsis.  
XX  
XX Example 1; Fig 1B1 to 1B-5; 47pp; English.  
XX  
XX This is the DNA sequence of plasmid pCDNA3 (Invitrogen). A portion of the

CC gene (see AAV38298) that codes for respiratory syncytial virus  
CC glycoprotein G (RSVG) has been amplified, digested with KpnI and ligated  
CC into KpnI-digested pcDNA3 upstream of the multiple cloning site of pcDNA3  
CC and downstream of the human cytomegalovirus immediate early (HCMV-IE)  
CC promoter to create pGT41. A full-length coding sequence of Streptococcus  
CC pneumoniae Rxi pneumococcal surface protein A (PspA) was then inserted  
CC into pGT41 to create a fusion between RSVG and PspA. Intramuscular  
CC immunisation of BALB/c mice with the resulting plasmid, designated  
CC pKSD2601, induced protection against an otherwise lethal challenge with a  
CC capsular type 3 pneumococcus. A plasmid for expression of pneumococcal  
CC epitope DNA in eukaryotic cells is claimed. The plasmid includes a  
CC promoter for driving expression in a eukaryotic cell (e.g. HCMV-IE), DNA  
CC encoding a leader sequence (e.g. of RSVG) which facilitates expression,  
CC translation through or transport of the expression product in a  
CC eukaryotic cell membrane, and DNA encoding a pneumococcal epitope such as  
CC PspA. The invention also provides a vaccine comprising the plasmid and a  
CC suitable carrier or diluent, and optionally one or more cytokines or DNA  
CC encoding them, or a bacterial delivery system. The vaccine is used to  
CC elicit an immunological response in a host, including humans, susceptible  
CC to pneumococcal infection or sepsis. The plasmid can also be used to  
CC express a pneumococcal epitope of interest in vitro. (Updated on 17-OCT-  
CC 2003 to standardise OS field)

XX SQ Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5446;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 60

DB 1204 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 1263

QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 101

DB 1264 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 1304

RESULT 14

AAS18619

ID AAS18619 standard; DNA; 5446 BP.

AC AAS18619;

XX 26-FEB-2002 (first entry)

DE Renilla luciferase vector pcDNA3.

XX Renilla luciferase; sea pansy; cell proliferation disorder;

KW immune disorder; hypogammaglobulinaemia; haematologic condition; anaemia;

KW neoplasm; cancer; human immunodeficiency virus; HIV;

KW tissue white cell infiltrative disorder; organ failure;

KW myotrophic condition; gonadal failure; bone disorder; muscle disorder;

KW osteoporosis; endocrine condition; vascular disorder; atherogenesis;

KW pcDNA3; ds.

XX Synthetic.

OS WO200181614-A2.

PN 01-NOV-2001.

XX 25-APR-2001; 2001WO-US013512.

PF 25-APR-2000; 2000US-00559874.

PR 02-JUN-2000; 2000US-00586339.

XX (CHEM-) CHEMICON INT.

PA Leng J;

XX WPI; 2002-041420/05.

DR

XX

PT

PT Determining cell proliferation for monitoring treatment of a subject,  
PT comprises obtaining light emission data from cell containing Renilla  
PT luciferase for specific time, and detecting a change in the data.

XX

PS Example 1; Fig 1A-B; 52pp; English.

XX

CC The invention describes a novel method for measuring proliferation of a  
CC cell or population of cells. The method comprises obtaining light  
CC emission data from a cell containing a Renilla luciferase over a period  
CC of time, cell proliferation of a cell or a population of cells can be  
CC measured by a change in light emission data indicating proliferation. A  
CC vector containing the Renilla luciferase enzyme is useful for diagnosing  
CC a cell proliferative disorder including: neoplasm or cancer, viral  
CC disorder or disease e.g. Human immunodeficiency virus (HIV), immune  
CC disorders e.g. hypogammaglobulinaemia and haematologic conditions e.g.  
CC anaemias, tissue white cell infiltrative disorders, organ failure,  
CC myotrophic conditions, gonadal failure, conditions of bone and muscle  
CC e.g. osteoporosis, endocrine conditions and vascular disorders e.g.  
CC atherogenesis, by transfecting a cell obtained from a subject with the  
CC vector and comparing the light emission data from the cell to that of a  
CC cell which does not have a cell proliferative disorder. A difference in  
CC light emission is indicative of a cell proliferative disorder. The vector  
CC is also useful for determining the effect of an agent on cell  
CC proliferation, by transfecting a cell obtained from a sample with the  
CC vector, and contacting the transfected cell with an agent suspected of  
CC modulating cell proliferation under conditions that allow the agent and  
CC the cell to interact, and comparing the light emission data from the cell  
CC to the light emission data from the cell in the absence of the agent. The  
CC methods are useful for drug discovery and drug screening, and in  
CC monitoring the treatment of a subject diagnosed with a cell proliferative  
CC disorder. This sequence is the vector pcDNA3 into which Renilla  
CC luciferase (AAS18616) is placed before transformation of cells with the  
CC vector to allow measurement of cell proliferation described in the method  
CC of the invention

XX

SQ Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 6; Length 5446;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 60

DB 1204 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATCGCGTGGCTCTA 1263

QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 101

DB 1264 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 1304

RESULT 15

ABL53540

ID ABL53540 standard; DNA; 5446 BP.

XX ABL53540;

XX 10-JUN-2002 (first entry)

DE Vector pcDNA3.

KW Vector; pcDNA3; heat shock protein 60; Hsp60; autoimmune disease;

KW insulin dependent diabetes mellitus; IDDM; DNA immunisation; vaccine;

KW CpG; antidiabetic; immunotherapy; gene therapy; ds.

OS Cytomegalovirus.

OS Bos taurus.

OS Unidentified.

OS Chimeric.

XX WO200216549-A2.

XX 28-FEB-2002.

XX





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 .Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_4141\_4241  
Perfect score: 101  
Sequence: 1 acagcaaggaggagattgg.....ccagctggggctctaggggg 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	295	CN778129	pgn2c.pk0
2	85	84.2	378	CF315931	HD--05-A1
3	70.6	69.9	640	CN161285	950563 MA
4	66.4	65.7	333	EX993205	Reverse s
5	66	65.3	772	BZ851581	CH240 281
6	62.6	62.0	755	CF366325	840970 MA
7	59.4	58.8	683	CF366325	840970 MA
8	58.6	58.0	648	CB419493	592356 MA
9	54.4	53.9	666	CR010518	Forward s
10	49.2	48.7	754	CR046705	Forward s
11	48.2	47.7	119	CR159589	Forward s
12	48.2	47.7	422	CB763949	AMGNNUC:M
13	48.2	47.7	433	CB760260	AMGNNUC:M
14	48.2	47.7	434	CB759281	AMGNNUC:M
15	48.2	47.7	449	CB742761	AMGNNUC:M
16	48.2	47.7	449	CB742762	AMGNNUC:M
17	48.2	47.7	450	CB788058	AMGNNUC:M
18	48.2	47.7	452	CB786684	AMGNNUC:M
19	48.2	47.7	457	CB740095	AMGNNUC:M
20	48.2	47.7	464	CB734940	AMGNNUC:M
21	48.2	47.7	467	CB714287	AMGNNUC:M
22	48.2	47.7	470	CB733007	AMGNNUC:M
23	48.2	47.7	472	CB730485	AMGNNUC:M
24	48.2	47.7	475	CB729030	AMGNNUC:M

25	48.2	47.7	482	6	CB728068	AMGNNUC:M
26	48.2	47.7	482	6	CB728069	AMGNNUC:M
27	48.2	47.7	482	6	CB728070	AMGNNUC:M
28	48.2	47.7	483	6	CB727859	AMGNNUC:M
29	48.2	47.7	484	6	CB727358	AMGNNUC:M
30	48.2	47.7	487	6	CB726627	AMGNNUC:M
31	48.2	47.7	487	6	CB726628	AMGNNUC:M
32	48.2	47.7	488	6	CB726362	AMGNNUC:M
33	48.2	47.7	496	6	CB713072	AMGNNUC:M
34	47.2	46.7	448	6	CB746136	AMGNNUC:M
35	47.2	46.7	450	6	CB788057	AMGNNUC:M
36	47.2	46.7	455	6	CB741728	AMGNNUC:M
37	46.2	45.7	439	6	CB750081	AMGNNUC:M
38	45	44.6	723	7	CK468080	939446 MA
39	44.6	44.2	207	9	CR162425	Forward s
40	43.4	43.0	798	7	CN158304	947089 MA
41	43.2	42.8	733	9	CR000527	Forward s
42	42.8	42.4	699	7	CN164031	994145 MA
43	42.8	42.4	778	7	CN161630	950941 MA
44	40.2	39.8	667	6	CB423182	596384 MA
45	39.6	39.2	752	7	CN164455	994805 MA

ALIGNMENTS

RESULT 1  
CN778129  
LOCUS CN778129 295 bp mRNA linear EST 20-MAY-2004  
DEFINITION pgn2c.pk001.h10.f Chicken Lymphoid cDNA library (pgn2c) Gallus  
gallus cDNA clone pgn2c.pk001.h10.f 3' end of pat.pk0008.d12 5',  
mRNA sequence.  
ACCESSION CN778129.1 GI:47548763  
VERSION EST.  
KEYWORDS Gallus gallus (chicken)  
SOURCE Gallus gallus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 295)  
AUTHORS Morgan, R.W. and Burnside, J.  
TITLE Chicken ESTs from lymphoid tissue- 3' sequence  
JOURNAL Unpublished (2004)  
COMMENT Contact: Robin W. Morgan  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1341  
Fax: 302-831-2822  
Email: morgan@udel.edu, www.chickest.udel.edu.  
FEATURES  
Location/Qualifiers  
source  
1..295  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="pgn2c.pk001.h10.f 3' end of pat.pk0008.d12"  
/sex="Male and Female"  
/tissue\_type="thymus, bursa, spleen, PBL, bone marrow"  
/lab\_host="E.Coli EMDH10B"  
/clone\_lib="Chicken Lymphoid cDNA library (pgn2c)"  
/note="Vector: pCMVSPORT 6"

ORIGIN  
Query Match 100.0%; Score 101; DB 7; Length 295;  
Best Local Similarity 100.0%; Pred. No. 9.8e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATACAGCATGCTGGGATCGCGTGGGCTCTA 60  
Db 188 ACAGCAAGGGGAGGATTGGGAAGACATACAGCATGCTGGGATCGCGTGGGCTCTA 247  
Qy 61 TGGCTTCTGAGCGGGAAGAACACAGCTGGGCTCTAGGGGG 101  
|||||

```

Db      248 TGGCTTCTGAGCGGAAGAACACCTGGGCTCTAGGGG 288

RESULT 2
LOCUS   CF315931
DEFINITION HD-05-A13.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD) Oryza sativa (japonica cultivar-group) cDNA clone HD-05-A13, mRNA sequence.
ACCESSION CF315931
VERSION   CF315931.1 GI:33687692
KEYWORDS EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 378)
AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
          Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University
          Yongin, Kyeonggi, Korea
          Tel: 82 31 330 6193
          Fax: 82 31 321 6355
          Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES             source
    source
    1..378
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nackdong"
    /db_xref="taxon:39947"
    /clone="HD-05-A13"
    /tissue_type="callus"
    /dev_stage="proliferated callus on 2N6 media for 2 weeks"
    /lab_host="E.coli DH10B"
    /clone_lib="OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD)"
    /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was treated with ABA(20um) for 1hr. Oligo-capped mRNA was reverse transcribed and then used for PCR. mRNA was derived from rice Histone Deacetylase overexpression line."

ORIGIN
Query Match      84.2%; Score 85; DB 7; Length 378;
Best Local Similarity 100.0%; Pred. No. 6.4e-15;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
        |||||
Db      293 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGGTGGGCTCTA 352
        |||||

Qy      61 TGGCTTCTGAGCGGAAGAACACG 85
        |||||
Db      353 TGGCTTCTGAGCGGAAGAACACG 377
        |||||

RESULT 3
LOCUS   CN161285
DEFINITION 950563 MARC 4P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CN161285
VERSION   CN161285.1 GI:46175715
KEYWORDS EST.
SOURCE   Sus scrofa (pig)
ORGANISM Sus scrofa
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 640)

AUTHORS Smith,T.P.L., Preking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329.
          Plate: TWM8058 row: C column: 15
          Seq primer: GTAATACGACTCACTATAGG.

FEATURES             source
    source
    1..640
    /organism="Sus scrofa"
    /mol_type="mRNA"
    /db_xref="taxon:9823"
    /tissue_type="pooled"
    /lab_host="DH10B"
    /clone_lib="MARC 4P1G"
    /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with combined RNA from day-10, day-13, day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match      69.9%; Score 70.6; DB 7; Length 640;
Best Local Similarity 81.2%; Pred. No. 1.5e-10;
Matches 82; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
        |||||
Db      499 ACACCAGGGGGGAGTGGGAACAATACCACTGGGGGAGCGGGGCTTTA 558
        |||||

Qy      61 TGGCTTCTGAGCGGAAGAACACGCTGGGGCTCTAGGGG 101
        |||||
Db      559 TGGCTTCTGAGCGGAAGAACACCTGGGGCTTTGGGGG 599
        |||||

RESULT 4
LOCUS   BX993205/c
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP272m10, genomic survey sequence.
ACCESSION BX993205
VERSION   BX993205.1 GI:49724663
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 333)
          Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
          Direct Submission
          Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES             source
    source
    1..333
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /clone="MHP272m10"
    /clone_lib="MHP"

ORIGIN
Query Match      65.7%; Score 66.4; DB 9; Length 333;
Best Local Similarity 87.5%; Pred. No. 2.4e-09;
Matches 84; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```

```

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
DB |||||
109 ACAGCAAGGGGAGGAGCTGGGAAGACAATAGCAGGCATGCTGGGAAAAA-GAGGGCTCTA 51
QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGCTCTA 96
DB |||||
50 TGGCTTCTGAGCGGGAAGAACCACTAGGCTTGTA 15

RESULT 5
BZ851581
LOCUS CH240_281G22.TJ CHORI-240 Bos taurus genomic clone CH240_281G22,
DEFINITION genomic survey sequence.
ACCESSION BZ851581
VERSION BZ851581.1 GI:29078986
KEYWORDS GSS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsengay,G., Geer,K.,
Shwartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
Crawford,A.M. and McEwan,J.C.
Bovine BAC End Sequences from Library CHORI-240
TITLE Unpublished (2003)
JOURNAL
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
Institute of Genomic Research (TIGR), USA.
Plate: 281 row: G column: 22
Seq primer: SP6
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..772
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_281G22"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Query Match 65.3%; Score 66; DB 8; Length 772;
Best Local Similarity 93.2%; Pred. No. 3.6e-09;
Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
DB |||||
219 ACAGCAAGGGGAGGATTGGGAGACATAGCAGGCATGCTGGGATCGGTGGGCTCTA 278
QY 61 TGGCTTCTGAGCGG 74
DB |||||
279 TGGGTACCCAGGTG 292

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
DB |||||
642 ACACCAAGGGGAGGATGGGAACCAATATCCGGCTGGCGGGGAGCGGGGGTTTTT 701
QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGCTCTAGGGGG 101
DB |||||
702 TGGTTTTTGTAGGGGAAAAAACCCCGTTGGGGCTTTAGGGGG 742

RESULT 7
CF366325
LOCUS CF366325
DEFINITION 840970 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CF366325
VERSION CF366325.1 GI:34169801
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 683)
Smith,T.P.L., Fekking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
A second set of porcine ESTs from a pooled-tissue normalized
library
TITLE Unpublished (2003)
JOURNAL

```

```

RESULT 6
CK462023
LOCUS CK462023
DEFINITION 932791 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CK462023
VERSION CK462023.1 GI:40833304
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 755)
Smith,T.P.L., Fekking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.E. and Keele,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
TITLE Unpublished (2003)
JOURNAL
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TW8036 row: C column: 13
Seq primer: GTAATACGACTCACTATAGG.
FEATURES
source
Location/Qualifiers
1..755
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 62.0%; Score 62.6; DB 7; Length 755;
Best Local Similarity 76.2%; Pred. No. 3.8e-08;
Matches 77; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
DB |||||
642 ACACCAAGGGGAGGATGGGAACCAATATCCGGCTGGCGGGGAGCGGGGGTTTTT 701
QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGCTCTAGGGGG 101
DB |||||
702 TGGTTTTTGTAGGGGAAAAAACCCCGTTGGGGCTTTAGGGGG 742

RESULT 7
CF366325
LOCUS CF366325
DEFINITION 840970 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CF366325
VERSION CF366325.1 GI:34169801
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 683)
Smith,T.P.L., Fekking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
A second set of porcine ESTs from a pooled-tissue normalized
library
TITLE Unpublished (2003)
JOURNAL

```

COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called with phred v0.020425.c and  
 trimmed with the aid of the trim\_alt option. Vector identified with  
 cross\_match v0.990329.  
 Plate: SRG8023 row: B column: 14  
 Seq primer: GTAATACGACTCACTATAGG.

# FEATURES

Location/Qualifiers  
 1..683  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="MARC 3P1G"  
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;  
 Library made with RNA pooled from multiple tissues  
 including brain, liver, muscle, placenta/endometrium,  
 ovary, testes, and bone marrow."

# ORIGIN

Query Match 58.8%; Score 59.4; DB 7; Length 683;  
 Best Local Similarity 74.3%; Pred. No. 3.4e-07;  
 Matches 75; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 ACACGAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTA 60  
 DB 526 ACACGAGGGGAGGATTGGGAAGAAAAAATAACAGCTTCCTGGGGATCCGGGGGCTTT 585  
 QY 61 TGGCTTCTGAGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG 101  
 DB 586 TGGTTCTGGGGGGAAGAACCCCTCGGGGCTTTGGGGG 626

RESULT 8  
 LOCUS CB419493 648 bp mRNA linear EST 25-MAR-2003  
 DEFINITION 592356 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION CB419493  
 VERSION CB419493.1 GI:29184608  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 648)  
 AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,  
 Wray,J.E. and Keele,J.W.  
 TITLE A second set of bovine ESTs from pooled-tissue normalized libraries  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called with phred v0.020425.c and  
 trimmed with the aid of the trim\_alt option. Vector identified with  
 cross\_match v0.990329.  
 Plate: FOY8013 row: P column: 9  
 Seq primer: GTAATACGACTCACTATAGG.

# FEATURES

Location/Qualifiers  
 1..648  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"

/clone\_lib="MARC 6BOV"  
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;  
 Library made with RNA pooled from multiple tissues  
 including liver, lung, hypothalamus, pituitary, and  
 placenta/endometrium."

# ORIGIN

Query Match 58.0%; Score 58.6; DB 6; Length 648;  
 Best Local Similarity 75.3%; Pred. No. 5.8e-07;  
 Matches 73; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 5 CACGGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTATGGC 64  
 DB 539 CCAGGGGGGAGGTTTGGAAAAACTTACCGGCTCTCTGGGGATCCGGGGGCTTTTGGT 598  
 QY 65 TTCTAGGGCGGAAAGAACCCAGCTGGGGCTCTAGGGGG 101  
 DB 599 TTTTGGCGGAAAAAACCACTTGGGGCTTTAGGGGG 635

# RESULT 9

LOCUS CR010518/c 666 bp DNA linear GSS 05-JUL-2004  
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and  
 chromosome engineering clone MHP232a16, genomic survey sequence.  
 ACCESSION CR010518  
 VERSION CR010518.1 GI:49743509  
 KEYWORDS GSS; genome survey sequence; MICER.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 666)  
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,  
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
 Rogers,J. and Bradley,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER

# FEATURES

Location/Qualifiers  
 1..666  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone\_lib="MHP232a16"  
 /clone\_lib="MHP232a16"

# ORIGIN

Query Match 53.9%; Score 54.4; DB 9; Length 666;  
 Best Local Similarity 72.9%; Pred. No. 1.1e-05;  
 Matches 70; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
 QY 2 CAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTAT 61  
 DB 110 CAGCAGGGGAGGATTAAGAGGACATAACAGCATGTTGGCAGGACGGTGGGTTTC 51  
 QY 62 GGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAG 97  
 DB 50 CGAGTATGCGTGGAAAGATCCAGCAGGGGCTGGAG 15

# RESULT 10

LOCUS CR046705/c 754 bp DNA linear GSS 05-JUL-2004  
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and  
 chromosome engineering clone MHP79m05, genomic survey sequence.  
 ACCESSION CR046705  
 VERSION CR046705.1 GI:49779760  
 KEYWORDS GSS; genome survey sequence; MICER.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 754)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>  
LOCATION/Qualifiers  
source  
1..754  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPP79m05"  
/clone\_lib="MHPP"

ORIGIN  
Query Match 48.7%; Score 49.2; DB 9; Length 754;  
Best Local Similarity 70.2%; Pred. No. 0.00039;  
Matches 66; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 3 AGCAAGGGGAGGATTTGGGAAGACATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATG 62  
|||||  
Db 108 AGCAGGGGAAGTATGCCAGACATAGCTGCTCCGCGTTGGGAAGGATCTTT 49  
|||||  
QY 63 GCTTCTAGGGGGAAGAACACAGCTGGGGCTCTA 96  
|||||  
Db 48 GCTTCTAGGGGGAAGAACCCAGGTGGCGATTTA 15  
|||||

RESULT 11  
CR159589/c  
LOCUS CR159589 119 bp DNA linear GSS 06-JUL-2004  
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and  
chromosome engineering clone MHP67p12, genomic survey sequence.  
ACCESSION CR159589  
VERSION CR159589.1 GI:49938438  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 119)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>  
LOCATION/Qualifiers  
source  
1..119  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP67p12"  
/clone\_lib="MHPP"

ORIGIN  
Query Match 47.7%; Score 48.2; DB 9; Length 119;  
Best Local Similarity 69.9%; Pred. No. 0.00056;  
Matches 65; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2 CAGCAAGGGGAGGATTTGGGAAGACATAGCAGGCATGCTGGGGATGCGGTGGGCTCTAT 61  
|||||  
Db 110 CTGCTTGGGGTGGTCTGGTGTCTTCTGCTGGCTGGCGTGGGCTCTTC 51  
|||||

QY 62 GCCTTCTAGGGGGAAGAACACAGCTGGGGCTC 94  
|||||  
Db 50 GGCCTCTGGGGTGTGTCTCTGCTGGGGCTC 18  
|||||

RESULT 12  
CB763949  
LOCUS CB763949 422 bp mRNA linear EST 16-MAY-2003

DEFINITION AMGNNUC:MRBE3-00121-F5-A rat brain E15 (10374) Rattus norvegicus  
CDNA clone mrbe3-00121-f5 5', mRNA sequence.  
ACCESSION CB763949  
VERSION CB763949.1 GI:29852340  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 422)  
AUTHORS Angen EST Program.  
TITLE Angen Rat EST Program  
JOURNAL Unpublished (2003)  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00121 row: f column: 5.  
LOCATION/Qualifiers  
source  
1..422

FEATURES  
source

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="mrbe3-00121-f5"  
/tissue\_type="brain E15"  
/clone\_lib="rat brain E15 (10374)"  
/note="Vector: pBCB; Site\_1: BatXI; Site\_2: NotI; rat  
brain E15"

ORIGIN

Query Match 47.7%; Score 48.2; DB 6; Length 422;  
Best Local Similarity 67.3%; Pred. No. 0.00071;  
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGAGGATTTGGGAAGACATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60  
|||||  
Db 286 ATAGTAAGGGGGAGGACCGCGAAGATACAGTAGTACGCGGAGCGTGGCGGCTCTCA 345  
|||||

QY 61 TGGCTTCTGAGCGGAAGACACAGCTGGGGCTCTAGGGGG 101  
|||||  
Db 346 CGGTCTCGAGGTGGAAGAATTAGTCGGGGCTCTGAGGGGG 386  
|||||

RESULT 13

CB760260  
LOCUS CB760260 433 bp mRNA linear EST 16-MAY-2003  
DEFINITION AMGNNUC:MRBE3-00121-G12-A rat brain E15 (10374) Rattus norvegicus  
CDNA clone mrbe3-00121-g12 5', mRNA sequence.

ACCESSION CB760260  
VERSION CB760260.1 GI:29848651  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 433)  
AUTHORS Angen EST Program.  
TITLE Angen Rat EST Program  
JOURNAL Unpublished (2003)  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00121 row: g column: 12.  
LOCATION/Qualifiers  
source  
1..433

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="mrbe3-00121-g12"  
/tissue\_type="brain E15"

```

/clone.lib="rat brain E15 (10374)"
/notes="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 433;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGTACGTGCGGACGTCGGCGGCTCTCA 345

Qy 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 CGGTCTCGAGGTGGAAGAATTAGTCGGGGTCTGAGGGGG 386

RESULT 14
CB759281
LOCUS
DEFINITION
AMGNNUC:MRBE3-00120-C9-A rat brain E15 (10374) Rattus norvegicus
cDNA clone mrbe3-00120-c9 5', mRNA sequence.
ACCESSION
CB759281
VERSION
CB759281.1 GI:29847672
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 434)
AUTHORS
Amgen EST Program.
TITLE
Amgen Rat EST Program
JOURNAL
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00120 row: c column: 9.
FEATURES
Location/Qualifiers
source
1..434
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrbe3-00120-C9"
/tissue_type="brain E15"
/clone.lib="rat brain E15 (10374)"
/notes="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 434;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGTACGTGCGGACGTCGGCGGCTCTCA 333

Qy 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CGGTCTCGAGGTGGAAGAATTAGTCGGGGTCTGAGGGGG 374

RESULT 15
CB742761
LOCUS
DEFINITION
AMGNNUC:MRBE3-00121-C7-A rat brain E15 (10374) Rattus norvegicus
cDNA clone mrbe3-00121-c7 5', mRNA sequence.
ACCESSION
CB742761
VERSION
CB742761.1 GI:29810059
KEYWORDS
EST.
```

```

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 449)
AUTHORS
Amgen EST Program.
TITLE
Amgen Rat EST Program
JOURNAL
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00121 row: c column: 7.
FEATURES
Location/Qualifiers
source
1..449
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrbe3-00121-C7"
/tissue_type="brain E15"
/clone.lib="rat brain E15 (10374)"
/notes="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 449;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGTACGTGCGGACGTCGGCGGCTCTCA 348

Qy 61 TGGCTTCTGAGCGGAAGAACACAGCTGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 CGGTCTCGAGGTGGAAGAATTAGTCGGGGTCTGAGGGGG 389

Search completed: July 14, 2005, 23:22:58
Job time : 968.667 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds

(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_8283\_8383

Perfect score: 101

Sequence: 1 aggggtattgtctcatgagc.....Gaaagtgcacctgacgtc 101

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	6 AR356490	AR356490 Sequence
C 2	101	100.0	142	6 AR338046	AR338046 Sequence
C 3	101	100.0	228	6 E00019	E00019 DNA coding
C 4	101	100.0	240	1 PMOENDO	M10199 Plasmid pM
C 5	101	100.0	251	6 E00018	E00018 DNA coding
C 6	101	100.0	251	6 I01644	I01644 Sequence 1
C 7	101	100.0	344	11 HUMUT5345	LI8624 Human chrom
C 8	101	100.0	400	6 BD195256	BD195256 Nucleotid
C 9	101	100.0	456	6 E00892	E00892 Synthetic D
C 10	101	100.0	456	6 E01156	E01156 DNA fragmen
C 11	101	100.0	456	6 E01274	E01274 DNA encodin
C 12	101	100.0	456	6 E01302	E01302 DNA encodin
C 13	101	100.0	466	6 AX260098	AX260098 Sequence
C 14	101	100.0	573	6 AX260150	AX260150 Sequence
C 15	101	100.0	693	6 A43586	A43586 Sequence 11
C 16	101	100.0	693	6 AR116755	AR116755 Sequence
C 17	101	100.0	998	1 AY559171	AY559171 Pseudomon
C 18	101	100.0	1011	1 SMTMAQGE	X97254 S.marcescen
C 19	101	100.0	1012	2 CEC11F10	Z92776 Caenorhabdi

20	101	100.0	1014	4 CFAJ4121	AJ224121 Canis fam
C 21	101	100.0	1027	1 AY589493	AY589493 Escherich
C 22	101	100.0	1040	1 AY538698	AY538698 Serratia
C 23	101	100.0	1040	1 AY538700	AY538700 Serratia
C 24	101	100.0	1040	1 AY538701	AY538701 Serratia
C 25	101	100.0	1040	1 AY538702	AY538702 Serratia
C 26	101	100.0	1041	1 AY538699	AY538699 Serratia
C 27	101	100.0	1042	1 AY394610	AJ394610 Klebsiell
C 28	101	100.0	1042	1 E0308558	AJ308558 Escherich
C 29	101	100.0	1044	1 AY392531	AJ392531 Streptoco
C 30	101	100.0	1044	1 AY452662	AF452662 Streptoco
C 31	101	100.0	1054	1 AF104441	AF104441 Klebsiell
C 32	101	100.0	1054	1 AF104442	AF104442 Escherich
C 33	101	100.0	1058	6 I03356	I03356 Sequence 4
C 34	101	100.0	1064	1 AY628199	AY628199 Escherich
C 35	101	100.0	1069	1 AF535127	AF535127 Klebsiell
C 36	101	100.0	1069	1 AY243512	AY243512 Klebsiell
C 37	101	100.0	1071	1 AY628175	AY628175 Escherich
C 38	101	100.0	1072	1 AY101764	AY101764 Klebsiell
C 39	101	100.0	1073	6 AR371489	AR371489 Sequence
C 40	101	100.0	1073	6 AX195443	AX195443 Sequence
C 41	101	100.0	1075	1 AY729027	AY729027 Proteus m
C 42	101	100.0	1075	1 PATN1PN2	X54606 Pseudomonas
C 43	101	100.0	1075	1 PATN2PN1B	X54607 Pseudomonas
C 44	101	100.0	1075	1 PATN3PN1A	X54604 Pseudomonas
C 45	101	100.0	1080	1 AF027199	AF027199 Klebsiell

ALIGNMENTS

RESULT 1  
LOCUS AR356490/c 142 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 2608 from patent US 6593114.  
ACCESSION AR356490  
VERSION AR356490.1 GI:33762574  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 142)  
AUTHORS Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.  
TITLE Staphylococcus aureus polynucleotides and sequences  
JOURNAL Patent: US 6593114-A 2608 15-JUL-2003;  
FEATURES Location/Qualifiers  
source 1..142  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 142;  
Best Local Similarity 100.0%; Pred. No. 8,7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAATAAACAAATAG 60  
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAATAAACAAATAG 48  
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101  
Db 47 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 7  
RESULT 2  
LOCUS AR538046/c 142 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 2608 from patent US 6737248.  
ACCESSION AR538046  
VERSION AR538046.1 GI:53929263  
KEYWORDS  
SOURCE Unknown.

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2408 18-MAY-2004;
FEATURES
    source
        Location/Qualifiers
            1..142
                /organism="unknown"
                /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 48
Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 228)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
    source
        Location/Qualifiers
            1..228
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /db_xref="taxon:562"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 116
Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101

```

```

Db 115 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75
RESULT 4
PMWENDO/c
LOCUS PMWENDO 240 bp DNA linear BCT 26-APR-1993
DEFINITION Plasmid pMW110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMW110
ORGANISM Plasmid pMW110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMW110 DNA.
FEATURES
    source
        Location/Qualifiers
            1..240
                /organism="Plasmid pMW110"
                /mol_type="genomic DNA"
                /db_xref="taxon:2599"
                /plasmid="Plasmid pMW110"
ORIGIN Unreported.
Query Match      100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAATAAG 92
Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 251)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;

```

EH Key Location/Qualifiers  
FH CDS 210..>252  
FT /product='E.coli penicillinase' FT  
TATA\_signal 190..196.  
FEATURES  
source Location/Qualifiers  
1..251  
/organism='Escherichia coli'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:562'  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 251;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 60  
|||||  
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 116  
|||||  
QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101  
|||||  
Db 115 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 75  
|||||  
RESULT 6  
I01644/c  
LOCUS 251 bp ss-DNA linear PAT 18-MAY-1993  
DEFINITION Sequence 1 from Patent US 4338397.  
ACCESSION I01644  
VERSION I01644.1 GI:267685  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 251)  
AUTHORS Gilbert,W. and Talmadge,K.  
TITLE Mature protein synthesis  
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;  
President and Fellows of Harvard College; Cambridge, MA  
FEATURES  
source Location/Qualifiers  
1..251  
/organism='unknown'  
/mol\_type='unassigned DNA'  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 251;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 60  
|||||  
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 116  
|||||  
QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101  
|||||  
Db 115 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 75  
|||||  
RESULT 7  
HUMUT5345  
LOCUS Human chromosome 8 STS UT5345, sequence tagged site.  
DEFINITION HUMUT5345 344 bp DNA linear STS 26-JUL-1993  
ACCESSION L18624  
VERSION L18624.1 GI:308338  
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;  
microsatellite repeat; repeat polymorphism; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 344)  
AUTHORS Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,  
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.  
Genetic and physical mapping of simple sequence repeat containing  
sequence tagged sites from the human genome  
Unpublished (1993)  
Original source text: Homo sapiens DNA.  
Submitted by: Utah Center for Human Genome Research University of  
Utah, Dept. of Human Genetics  
2160 Eccles Institute of Human Genetics  
Salt Lake City, UT 84112  
e-mail: sts@corona.med.utah.edu  
Primer A: GAGCAAAAACAGGAGGCAAAATGC  
Primer B: TTCGGGAAATGTGCGCGAACC  
32P-label: B Primer  
PCR Profile:  
Initial Denaturation: 94C 300sec  
PCR Cycles: 30  
Denaturation: 94C 10sec  
Annealing: 60C 10sec  
Extension: 72C 20sec  
Mg++: 2mM  
Gel: Acrylamide 7%, Formamide 32%, Urea 34%  
Alleles: 2  
FEATURES  
source Location/Qualifiers  
1..344  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'  
/map='8'  
36..224  
/standard\_name='STS UT5345'  
36..60  
primer\_bind  
primer\_bind complement (202..224)  
ORIGIN  
Query Match 100.0%; Score 101; DB 11; Length 344;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 60  
|||||  
Db 141 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAAG 200  
|||||  
QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101  
|||||  
Db 201 GGGTTCCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 241  
|||||  
RESULT 8  
BD195256/c  
LOCUS BD195256 400 bp DNA linear PAT 17-JUL-2003  
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.  
ACCESSION BD195256  
VERSION BD195256.1 GI:33005021  
KEYWORDS JP 2002513277-A/43.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE  
1 (bases 1 to 400)  
AUTHORS Dillon,P.J., Choi,G.H. and Welch,R.A.  
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands  
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;  
HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION  
COMMENT OS Unidentified  
PN JP 2002513277-A/43  
PD 08-MAY-2002  
PF 21-NOV-1997 JP 1998523916  
PR 22-NOV-1996 US 60/031626.14-OCT-1997 US 60/061953 PI  
PATRICK J DILLON,GIL H CHOI,RODNEY A WELCH  
PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC  
Strandedness: Double;  
CC Topology: Linear;  
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
FEATURES
source 1..400
/organism='Unidentified'.
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 60
Db 165 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 106

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101
Db 105 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 65

RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai, H., Momota, Y., Kumakura, T., Tochifusa, N., Kitazawa, T.,
Ojida, K. and Matsushiro, A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
TOCHIFUSA NORIYUKI,
KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO
PC C12N15/00, C12N1/20, C12P21/00, (C12N1/20, C12R1:19), (C12P21/00, PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clone=pVG201;
CC Feature is identified by experimental;
FH key Location/Qualifiers
FH promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
FEATURES
source 1..456
/organism='synthetic construct'
/mol_type='genomic DNA'

```

```

ORIGIN
/db_xref="taxon:32630"

Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 73

RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa, K., Momota, Y., Kajifusa, N., Koide, T. and Okai, H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
OKAI HIDEO
PC C12N15/00, C12N1/20, C12P21/00, (C12N1/20, C12R1:19), (C12N1/20, PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH key Location/Qualifiers
FH promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FT Location/Qualifiers
FEATURES
source 1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101

```

```

1. .466
/organism="Drosophila melanogaster"
/mol type="unassigned DNA"

```

/db\_xref="taxon:7227"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 466;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60  
|||||  
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 221  
|||||

Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 101  
|||||  
Db 220 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 180  
|||||

## RESULT 14

AX260150/c AX260150 573 bp DNA linear PAT 26-OCT-2001  
LOCUS  
DEFINITION Sequence 112 from Patent WO0172774.  
ACCESSION AX260150  
VERSION AX260150.1 GI:16509172

## KEYWORDS

SOURCE Drosophila melanogaster (fruit fly)

## ORGANISM

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1  
Deak, P., Glover, D.M. and Midgley, C.  
Cell cycle progression proteins  
Patent: WO 0172774-A 112 04-OCT-2001;  
Cyclacel Limited (GB)

## FEATURES

source  
1..573  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:7227"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 573;  
Best Local Similarity 100.0%; Pred. No. 8.8e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60  
|||||  
Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 296  
|||||

Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 101  
|||||  
Db 295 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 255  
|||||

## RESULT 15

A43586 A43586 693 bp DNA linear PAT 06-MAR-1997  
LOCUS  
DEFINITION Sequence 11 from Patent WO9507357.  
ACCESSION A43586  
VERSION A43586.1 GI:2298779

## KEYWORDS

SOURCE Cuphea lanceolata

## ORGANISM

Cuphea lanceolata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Myrtales; Lythraceae; Cuphea.

## REFERENCE

1 (bases 1 to 693)  
Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,  
Hoerliche-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,  
Schulte, W., Voetz, M., Walek, J. and Schell, J.

## AUTHORS

## PROMOTERS

## TITLE

## JOURNAL

## COMMENT

Patent: WO 9507357-A 11 16-MAR-1995;  
MAX PLANCK GESELLSCHAFT (DE)  
Other publication CA 2169093 950316

Other publication AU 7615494 950327.

## FEATURES

source  
1..693  
Location/Qualifiers  
/organism="Cuphea lanceolata"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3930"  
/clone="CLKASIG8"  
/clone\_lib="Genomic Lambda Fix II"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 693;  
Best Local Similarity 100.0%; Pred. No. 8.8e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60  
|||||  
Db 592 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 651  
|||||

Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 101  
|||||  
Db 652 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 692  
|||||

Search completed: July 14, 2005, 14:03:30

Job time : 756.618 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_8283\_8383  
Perfect score: 101  
Sequence: 1 aggtttattgtctcatgagc.....gaaagtgcacacctgacgtc 101

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_16Dec04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2002bs:\*
  - 8: Geneseqn2003as:\*
  - 9: Geneseqn2003bs:\*
  - 10: Geneseqn2003cs:\*
  - 11: Geneseqn2003ds:\*
  - 12: Geneseqn2004as:\*
  - 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2	AAV76919
C 2	101	100.0	228	1	AAV76919 Staphyloc
C 3	101	100.0	251	1	Aan10032 Sequence
C 4	101	100.0	400	2	Aan10031 Sequence
C 5	101	100.0	456	1	AAV31229 E. coli J
C 6	101	100.0	456	1	Aan60824 Plasmid p
C 7	101	100.0	456	1	Aan71080 Sequence
C 8	101	100.0	456	1	Aan70833 Beta-urog
C 9	101	100.0	466	6	Aan81765 Sequence
C 10	101	100.0	487	2	AbA90413 Drosophil
C 11	101	100.0	535	2	Aax21173 Polynucle
C 12	101	100.0	573	6	Aax21149 Polynucle
C 13	101	100.0	605	12	AbA90456 Drosophil
C 14	101	100.0	776	4	Adh58311 Electroph
C 15	101	100.0	776	4	Aas30560 DNA encod
C 16	101	100.0	776	4	Aas27819 DNA encod
C 17	101	100.0	776	4	Abx42984 Genomic s
C 18	101	100.0	776	4	Aal07344 Human rep
C 19	101	100.0	776	4	Aal03229 Human rep
C 20	101	100.0	776	4	Aal06588 Human rep
C 21	101	100.0	776	4	Aal07340 Human rep

C 21	101	100.0	776	5	ABA14573
C 22	101	100.0	776	5	AAS34681
C 23	101	100.0	776	8	ADA41574
C 24	101	100.0	776	8	ACC50905
C 25	101	100.0	776	8	ABZ71508
C 26	101	100.0	776	9	ADB91869
C 27	101	100.0	776	9	ADB61140
C 28	101	100.0	776	10	ADB94622
C 29	101	100.0	776	10	ADC74663
C 30	101	100.0	776	10	ADA57709
C 31	101	100.0	776	12	ADN41551
C 32	101	100.0	845	4	AAS30559
C 33	101	100.0	845	4	AAS27818
C 34	101	100.0	845	4	ABK42983
C 35	101	100.0	845	4	AAS41807
C 36	101	100.0	845	4	AAK41855
C 37	101	100.0	845	4	AAK85485
C 38	101	100.0	845	4	AAK85434
C 39	101	100.0	845	4	AAAL07343
C 40	101	100.0	845	4	AAAL06587
C 41	101	100.0	845	4	AAAL07339
C 42	101	100.0	845	4	AAAL03228
C 43	101	100.0	845	5	ABA14572
C 44	101	100.0	845	5	AAS34680
C 45	101	100.0	845	9	ADB61139

ALIGNMENTS

RESULT 1  
AAV76919/c  
ID AAV76919 standard; DNA; 142 BP.

XX AC AAV76919;

XX DT 16-MAR-1999 (first entry)

XX STaphylococcus aureus contig SEQ ID #2608.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.

XX OS Staphylococcus aureus.

XX EP786519-A2.

XX PD 30-JUL-1997.

XX PF 07-JAN-1997; 97EP-00100117.

XX PR 05-JAN-1996; 96US-0009861P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
stored on computer readable medium and used in the production of anti-  
S.aureus vaccines.

XX Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
of the invention. The DNA sequences are recorded on a computer readable  
medium, preferably selected from a floppy or hard disk, random access  
memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
the S.aureus DNA sequences allows putative functions to be assigned so  
that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against *S. aureus* infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the *S. aureus* DNA sequences contained on the computer  
CC readable medium  
XX  
SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.1e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 AGGGTTATGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60  
Db 107 AGGGTTATGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 48  
Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101  
Db 47 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 7

RESULT 2  
AA10032/c  
ID AA10032 standard; DNA; 228 BP.  
XX  
AC AA10032;  
XX  
DT 13-AUG-1992 (first entry)  
XX  
DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.  
XX  
KW Cloning vehicle; bacterial vector; transformed host; penicillinase;  
KW insulin; ds.  
XX  
OS Escherichia coli.

FH Key Location/Qualifiers  
FT misc\_feature 1..4  
FT /\*tag= a  
FT /label= sticky end  
FT misc\_feature 225..228  
FT /\*tag= b  
FT /label= sticky end  
XX  
FN EP38182-A.  
XX  
PD 21-OCT-1981.  
XX  
PF 09-APR-1981; 81EP-00301561.  
XX  
PR 11-APR-1980; 80US-00139225.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Gilbert W, Talmadge K;  
XX  
DR WPI; 1981-80125D/44.  
DR P-PSDB; AAP10039.  
XX  
PT Synthesis of mature protein or polypeptide - by using bacterial host  
PT transformed by cloned vehicle contg. DNA fragment etc.  
XX  
PS Example; Fig 3; 34pp; English.  
XX  
CC The closest identifiable promoter for the penicillinase gene in pKT241  
CC (AA10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was  
CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
CC preproinsulin (see AA10033). The closest identifiable promoter for the  
CC penicillinase gene in pKT218 (AA10032) is located in the region 14 to 20  
CC nucleotides before its translational start signal. In the examples, the  
CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
CC fragment (CB6) for rat preproinsulin (see AA10034)  
XX

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60  
Db 175 AGGGTTATGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 116  
Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101  
Db 115 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75

RESULT 3  
AA10031/c  
ID AA10031 standard; DNA; 251 BP.  
XX  
AC AA10031;  
XX  
DT 13-AUG-1992 (first entry)  
XX  
DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.  
XX  
KW Cloning vehicle; bacterial vector; transformed host; penicillinase;  
KW insulin; ds.  
XX  
OS Escherichia coli.

FH Key Location/Qualifiers  
FT misc\_feature 1..4  
FT /\*tag= a  
FT /label= sticky end  
FT misc\_feature 248..251  
FT /\*tag= b  
FT /label= sticky end  
XX  
FN EP38182-A.  
XX  
PD 21-OCT-1981.  
XX  
PF 09-APR-1981; 81EP-00301561.  
XX  
PR 11-APR-1980; 80US-00139225.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Gilbert W, Talmadge K;  
XX  
DR WPI; 1981-80125D/44.  
DR P-PSDB; AAP10038.  
XX  
PT Synthesis of mature protein or polypeptide - by using bacterial host  
PT transformed by cloned vehicle contg. DNA fragment etc.  
XX  
PS Example; Fig 2; 34pp; English.

XX  
CC The closest identifiable promoter for the penicillinase gene in pKT241  
CC (AA10031) is located in the region 14 to 20 nucleotides before its  
CC translational start signal. In the examples, the 3' end of pKT241 was  
CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
CC preproinsulin (see AA10033). The closest identifiable promoter for the  
CC penicillinase gene in pKT218 (AA10032) is located in the region 14 to 20  
CC nucleotides before its translational start signal. In the examples, the



```
CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)
XX
SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
Db |||||
175 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 116
QY 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db |||||
115 GGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTC 75

RESULT 4
AAV31229/c
ID AAV31229 standard; DNA; 400 BP.
XX
AC AAV31229;
XX
DT 01-OCT-1998 (first entry)
XX
DE E. coli J96 pathogenicity island contig #43.
XX
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW PAI V; pheV; vaccine; protective immune response; ds.
XX
OS Escherichia coli.
XX
PN WO9822575-A2.
XX
PD 28-MAY-1998.
XX
PF 21-NOV-1997; 97WO-US021347.
XX
PR 22-NOV-1996; 96US-0031626P.
PR 14-OCT-1997; 97US-0061953P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (UYWI-) UNIV WISCONSIN.
XX
XX Dillon PJ, Choi GH, Welch RA;
XX WPI; 1998-312461/27.
XX
New isolated uropathogenic E. coli nucleotide sequences - used to develop
products for the detection of pathogenic E. coli and to elicit an immune
response to pathogenic E. coli.
XX
Claim 21; Page 140-141; 250pp; English.
XX
This sequence represents a E. coli strain J96 contig containing
pathogenicity island (PAI) sequences, and represents a nucleic acid
molecule of the invention. PAIs are large fragments of DNA which comprise
pathogenicity determinants. The sequences of the invention are taken from
PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
on the E. coli chromosome and is greater than 170 kb. PAI V is located at
approximately 94 min (at pheR) on the E. coli chromosome and is
approximately 160 kb in size. Antibodies specific to the proteins encoded
by the PAI open reading frames of the invention can be used in kits to
detect uropathogenic E. coli. The proteins are used in vaccines to elicit
a protective immune response in an animal to the uropathogenic E. coli
strain J96
XX
SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
Db |||||
165 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 106
QY 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db |||||
105 GGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTC 65

RESULT 5
AAN60624/c
ID AAN60624 standard; DNA; 456 BP.
XX
AC AAN60624;
XX
DT 25-MAR-2003 (revised)
DT 29-OCT-1991 (first entry)
XX
DE Plasmid pUG201 sequence encoding beta-urogastrone.
XX
KW Beta-lactamase signal peptide; pGH54; pGH55; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT promoter 125..170
FT /*tag= a
FT RBS 200..203
FT /*tag= b
FT CDS 209..439
FT /*tag= c
FT sig_peptide 209..277
FT /*tag= d
FT /*label= Beta-lactamase signal peptide
FT mat_peptide 278..436
FT /*tag= e
FT /*label= Beta-urogastrone
XX
WO8603779-A.
XX
PN 03-JUL-1986.
XX
PF 19-DEC-1985; 85WO-JP000696.
XX
PR 21-DEC-1984; 84JP-00271206.
XX
PA (EARTH) EARTH CHEM CO LTD.
PA (OHGA/) OHGAI H.
XX
PI Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;
XX WPI; 1986-182911/28.
XX P-PSDB; AAP60678.
XX
Recombinant vector for polypeptide secretion - contains signal peptide
sequence directly bonded to peptide-coding sequence.
XX
Disclosure; Table 4; 79pp; Japanese.
XX
The plasmid produces secreted beta-urogastrone in a transformed
expression system. Similar plasmids may be constructed where the
secretion signal may be coupled with eg. somatostatin, insulin, growth
hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 6
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 7
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)

```

```

AAN70833/c
ID AAN70833 standard; DNA; 456 BP.
XX
AC AAN70833;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
XX Beta-urogastrone sequence.
XX
XX Tumour; inosine; DNA probe; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= b
FT RBS 200..204
FT /*tag= c
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
DR P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
PT using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
CC polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
CC ssDNA and probe are hybridized and the existence of DNA in the product is
CC detected. It can be used to detect the presence of malignant tumour.
CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)

```

```
XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 209..277 /*tag= a
XX FT 278..439 /*tag= b
XX FT /*product= "New beta-urogastrone deriv."
XX
XX JP63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
XX
XX P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
XX secretion inhibiting action, or cell proliferation promoting action. The
XX deriv. has the same biological or pharmacological activities as beta-
XX urogastrone. It is not susceptible to denaturation by oxidn. and is
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX pepsinase. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
XX 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 114
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73
XX
XX RESULT 9
XX ABA90413/c
XX ID ABA90413 standard; DNA; 466 BP.
XX
XX AC ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antiinflammatory; antiposrotatic; dermatological; antifungal; mitosis;
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
XX
XX WO200172774-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-GB001297.
XX
XX 24-MAR-2000; 2000GB-00007268.
XX
XX (CYCL-) CYCLACEL LTD.
XX
XX Deak P, Glover DM, Midgley C;
XX
XX WPI; 2002-055132/07.
XX
XX Polynucleotides encoding cell cycle progression proteins, useful for
XX treating a tumor or a proliferative disorder.
XX
XX Claim 1; Page 99; 213pp; English.
XX
XX The present invention relates to Drosophila cell cycle progression
XX proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
XX ABA90520). The coding sequences and proteins are useful for identifying a
XX substance capable of affecting the function of the corresponding gene, a
XX inhibiting mitosis and/or meiosis. They can also be used in a method for
XX treating a tumour or proliferative disorder, cardiovascular disorders
XX (such as restenosis and cardiomyopathy), autoimmune disorders such as
XX glomerulonephritis and rheumatoid arthritis), dermatological disorders
XX (such as psoriasis), antiinflammatory, antifungal and antiparasitic
XX disorders (such as malaria)
XX
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 101; DB 6; Length 466;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
XX 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 221
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
XX
XX RESULT 10
XX AAX21173/c
XX ID AAX21173 standard; DNA; 487 BP.
XX
XX AC AAX21173;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX
```

DR WPI; 1999-081273/07.

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1106; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60

DB 323 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 264

QY 61 GGGTCCCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

DB 263 GGGTCCCGCACATTTCCCGAAAAGTGCACCTGACGTC 223

RESULT 11

AAX21149/C

ID AAX21149 standard; DNA; 535 BP.

XX

AC AAX21149;

XX

XX 05-MAY-1999 (first entry)

XX

DE Polynucleotide sequence from the genome of *Treponema pallidum*.

XX

DE *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;

KW enzyme production; ds.

XX

OS *Treponema pallidum*.

XX

XX WO9859034-A2.

XX

PN 30-DEC-1998.

XX

XX 23-JUN-1998; 98WO-US013041.

XX

XX 24-JUN-1997; 97US-0050667P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Fraser CM;

XX

DR WPI; 1999-081273/07.

XX

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1093; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60

DB 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 296

Query Match 100.0%; Score 101; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60

DB 158 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 99

QY 61 GGGTCCCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

DB 98 GGGTCCCGCACATTTCCCGAAAAGTGCACCTGACGTC 58

RESULT 12

ABA90456/C

ID ABA90456 standard; DNA; 573 BP.

XX

AC ABA90456;

XX

XX 12-FEB-2002 (first entry)

XX

DE *Drosophila* cell cycle progression protein coding sequence #31.

XX

KW Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;

KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;

KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;

KW cell cycle progression protein; tumour; proliferative disorder;

KW cardiovascular; autoimmune; dermatological disorder; ds.

XX

OS *Drosophila* sp.

XX

XX WO200172774-A2.

XX

XX 04-OCT-2001.

XX

XX 23-MAR-2001; 2001WO-GB001297.

XX

XX 24-MAR-2000; 2000GB-00007268.

XX

PA (CYCL-) CYCLACEL LTD.

XX

PI Deak P, Glover DM, Midgley C;

XX

XX WPI; 2002-055132/07.

XX

XX Polynucleotides encoding cell cycle progression proteins, useful for

PT treating a tumor or a proliferative disorder.

XX

PS Claim 1; Page 144; 213pp; English.

XX

CC The present invention relates to *Drosophila* cell cycle progression

CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-

CC ABA90520). The coding sequences and proteins are useful for identifying a

CC substance capable of affecting the function of the corresponding gene, a

CC substance capable of inhibiting the cell division cycle, or capable of

CC inhibiting mitosis and/or meiosis. They can also be used in a method for

CC treating a tumour or proliferative disorder, cardiovascular disorders

CC (such as restenosis and cardiomyopathy), autoimmune disorders such as

CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders

CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic

CC disorders (such as malaria)

XX

SQ Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60

DB 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 296

QY 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101  
Db 295 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 255

RESULT 13

ADH58311  
ID ADH58311 standard; DNA; 605 BP.

AC ADH58311;

XX 25-MAR-2004 (first entry)

DT Electropherogram of a DNA sequencing reaction using E154 & T422.

DE ds; primer library; extendable oligos; EO; ligation chain reaction; LCR;  
KW rolling circle amplification; strand displacement amplification;  
KW isothermal DNA amplification; biotechnology; agriculture;  
KW medical research; pUC19 plasmid.

XX Synthetic.

OS Escherichia coli.

XX WO2003093500-A1.

XX 13-NOV-2003.

XX 24-DEC-2002; 2002WO-AU001763.

XX 01-MAY-2002; 2002AU-00002045.

XX (NUCL-) NUCLEICS PTY LTD.

XX Tillett D, Thomas T;

XX WPI; 2004-053046/05.

XX Increasing the affinity of an extendable oligonucleotide (EO) for a  
PT target nucleic acid, for providing primers having improved specificity,  
PT comprises hybridization of the EO to a template oligonucleotide (TO) and  
PT extension of the EO.

PS Example 10; Fig 23; 85pp; English.

XX This invention relates to a novel method for the optimisation of primer  
CC libraries. Specifically, it refers to increasing the affinity of short  
CC oligonucleotide primers, also known as extendable oligos (EOs), for their  
CC template sequences. The present invention describes improved methods for  
CC sequencing and the linear and exponential amplification of DNA that can  
CC be useful for PCR, RT-PCR, ligation chain reaction (LCR), rolling circle  
CC amplification, strand displacement amplification and isothermal DNA  
CC amplification. Accordingly, these extendable oligos with improved  
CC specificity and affinity are particularly important in fields ranging  
CC from biotechnology and agriculture to medical research. This  
CC polynucleotide sequence is the electropherogram of a DNA sequencing  
CC reaction that used the pUC19 plasmid and E154/T422 oligos, used in an  
CC exemplification of the invention.

XX Sequence 605 BP; 159 A; 133 C; 147 G; 148 T; 0 U; 18 Other;

Query Match 100.0%; Score 101; DB 12; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.8e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGACGGATACATATTTGAATGTTAGAAAAATAACAATAAG 60

Db 259 AGGGTTATTGTCATGACGGATACATATTTGAATGTTAGAAAAATAACAATAAG 318

QY 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101

Db 319 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 359

RESULT 14

AAS30560/c

ID AAS30560 standard; DNA; 776 BP.

XX AC AAS30560;

XX 21-NOV-2001 (first entry)

DT DNA encoding novel prostate gland antigen, Seq ID No 418.

DE Human; neutrotropic; neuroprotective; cytostatic; antiparkinsonian;  
KW antineoplastic; dermatological; immunosuppressive; antiinflammatory;  
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;  
KW osteoarthritic; prostate gland; prostatitis; adenocarcinoma; hair loss;  
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;  
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;  
KW reproductive system disorder; autoimmune disorder; urinary system;  
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;  
KW blood-related disorder; hyperproliferative disorder; respiratory;  
KW neurological disorder; endocrine disorder; inflammatory disorder;  
KW liver disorder; wound healing; food preservative; ds.

XX Homo sapiens.

XX WO200155447-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001330.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226868P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 06-SEP-2000; 2000US-0230437P.



XX AAS27819;  
XX 07-NOV-2001 (first entry)  
XX DNA encoding novel signal transduction pathway protein, Seq ID 1479.  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX WO200154733-A1.  
XX 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001312.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241836P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.

PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254037P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465460/50.	
XX		
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or	
PT	prognosing disorders related to the proteins, including cancers, immune	
PT	disorders and neuronal disorders.	
XX		
PS	Claim 1; SEQ ID NO 1479; 880pp; English.	
XX		
CC	The invention relates to novel isolated polypeptides (I), and	
CC	polynucleotides (II). (I), (II) and the antibody to (I) are useful for	
CC	diagnosing, preventing and treating diseases including immune system	
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune	
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ	
CC	transplant rejections and graft versus host disease, infectious diseases	
CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and	
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative	
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders	
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities	
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.	
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),	
CC	respiratory disorders, dermatological disorders, in wound healing,	
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's	
CC	disease), reproductive system disorders, gastrointestinal disorder	
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of	
CC	B-cell responsiveness to pathogens, activators of T-cells, to induce	
CC	higher affinity antibodies, and as a means to induce tumour proliferation	
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-	
CC	AAS27850 represent novel signal transduction pathway protein coding	
CC	sequences and PCR primers of the invention	
XX		
SQ	Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;	
	Query Match 100.0%; Score 101; DB 4; Length 776;	
	Best Local Similarity 100.0%; Pred. No. 2.9e-21;	
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Oy	1 AGGCTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60	
Dd	546 AGGGTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 487	
Oy	61 GGGTTCGCGCACATTTCCCCGAAAAGTGCCACTGCAGTC 101	
Dd	486 GGGTTCGCGCACATTTCCCCGAAAAGTGCCACTGCAGTC 446	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-43\_COPY\_8283\_8383  
Perfect score: 101  
Sequence: 1 agggttattgtctcatgagc.....gaaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095 83374 Heb
C 2	101	100.0	300	5	BU963956 EST88 Heb
C 3	101	100.0	300	5	BU964094 EST226 He
C 4	101	100.0	309	9	AL000426 F.rubripe
C 5	101	100.0	391	1	AL597149 DKEZp313J
C 6	101	100.0	414	9	CC819240 100005D19
C 7	101	100.0	417	4	BJ684174 BJ684174
C 8	101	100.0	491	9	CC811923 100006J13
C 9	101	100.0	495	4	B1805285 S035A01 S
C 10	101	100.0	495	9	CC818374 100004B07
C 11	101	100.0	496	9	CC818523 100004L13
C 12	101	100.0	503	9	CC8119854 100006N08
C 13	101	100.0	515	9	CC8117752 100003C16
C 14	101	100.0	518	9	CC8117128 100002D21
C 15	101	100.0	519	9	CC8117162 100002J19
C 16	101	100.0	519	9	CC8117796 100003K14
C 17	101	100.0	521	9	CC8119067 100005C09
C 18	101	100.0	533	9	CC8119841 100006L07
C 19	101	100.0	542	9	CC8116892 100002L01
C 20	101	100.0	550	7	CB766522 DKEZp469H
C 21	101	100.0	551	9	CC8116905 100002N02
C 22	101	100.0	554	9	CC8119058 100005A09
C 23	101	100.0	563	9	CC8119270 100005G21
C 24	101	100.0	566	9	CC8116848 100002D02

C 25	101	100.0	566	9	CC820024 100006K23
C 26	101	100.0	567	9	CC817070 100002J15
C 27	101	100.0	568	9	CC818640 100004P22
C 28	101	100.0	571	9	CC816954 100002F11
C 29	101	100.0	571	9	CC818423 100004J12
C 30	101	100.0	580	9	CC819098 100005I07
C 31	101	100.0	582	1	AL694813 DKEZp313I
C 32	101	100.0	583	9	CC817633 100003M06
C 33	101	100.0	583	9	CC818436 100004M08
C 34	101	100.0	586	9	CC816883 100002J03
C 35	101	100.0	588	9	CC817788 100003I18
C 36	101	100.0	588	9	CC818340 100004K04
C 37	101	100.0	589	9	CC817595 100003S03
C 38	101	100.0	590	9	CC819754 100006L06
C 39	101	100.0	592	9	CC817679 100003F10
C 40	101	100.0	592	9	CC818508 100004I16
C 41	101	100.0	593	9	CC816942 100002D10
C 42	101	100.0	593	9	CC817699 100003J09
C 43	101	100.0	594	9	CC818287 100004A02
C 44	101	100.0	594	9	CC818422 100004J11
C 45	101	100.0	595	9	CC816929 100002B08

ALIGNMENTS

RESULT 1  
BM078095/c 300 bp mRNA linear EST 30-NOV-2001  
LOCUS 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma  
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma  
ACCESSION BM078095  
VERSION BM078095.1 GI:17157967  
KEYWORDS EST.  
SOURCE Hebeloma cylindrosporum  
ORGANISM Hebeloma cylindrosporum  
REFERENCE 1 (bases 1 to 300)  
AUTHORS Wipf D., Benjdia M., Tegeder M. and Frommer W.B.  
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wipf D  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: pDR196 5' primer (PMA 5')  
HIGH quality sequence stop: 300  
POLYA-No.

FEATURES  
source  
1..300  
/organism="Hebeloma cylindrosporum"  
/mol\_type="mRNA"  
/strain="H1"  
/db\_xref="taxon:76867"  
/tissue\_type="Mycelia"  
/lab\_host="E. coli XLI-Blue"  
/clone\_lib="Hebeloma cylindrosporum functional cDNA library"  
/note="vector: pDR 196 (unpublished); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 8.1e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60

```

|||||
174 AGGGTTATGTCATGACGGATACATATTGAATGATTAGAAAAATAACAAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c 300 bp mRNA linear EST 13-NOV-2002
LOCUS ESP88 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycetes; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .

FEATURES
source
Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_hosts="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGACGGATACATATTGAATGATTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATGTCATGACGGATACATATTGAATGATTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140 309 bp DNA linear GSS 25-FEB-2004
LOCUS F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.

FEATURES
source
Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_hosts="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGACGGATACATATTGAATGATTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATGTCATGACGGATACATATTGAATGATTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140 309 bp DNA linear GSS 25-FEB-2004
LOCUS F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 98
Qy 61 GGGTTCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101
Db 99 GGGTTCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 139
RESULT 5
AL597149
LOCUS
DEFINITION
DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
AL597149
ACCESSION
VERSION
KEYWORDS
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
REFERENCE
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
    Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
TITLE
    EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL
    Unpublished (1999)
COMMENT
    Contact: MIPS
    MIPS
    Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
    This is the 5' sequence of the clone insert
    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
    sequenced by BMF2 (Biomedical Research Center at the Charite,
    Berlin/Germany) within the cDNA sequencing consortium of the German
    Genome Project.
    No sl sequence available.
    This clone (DKFZp313J1611) is available at the RZPD in Berlin.
    Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
    Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stages="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: SfIra; Site_2: SfiIB;
                cDNA-collection"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 98
Qy 61 GGGTTCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101
Db 99 GGGTTCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 139
RESULT 6
CC819240
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
    Sterkiella histriomuscorum
    Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
    Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
AUTHORS
    Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
JOURNAL
    Unpublished (2003)
COMMENT
    Contact: Robert B. Weiss
    University of Utah Genome Center
    University of Utah
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC100005D19"
                /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 355
RESULT 6
CC819240/c
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
    Sterkiella histriomuscorum
    Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
    Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
AUTHORS
    Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
JOURNAL
    Unpublished (2003)
COMMENT
    Contact: Robert B. Weiss
    University of Utah Genome Center
    University of Utah
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC100005D19"
                /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 355
RESULT 6
CC819240/c
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
    Sterkiella histriomuscorum
    Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
    Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
AUTHORS
    Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
JOURNAL
    Unpublished (2003)
COMMENT
    Contact: Robert B. Weiss
    University of Utah Genome Center
    University of Utah
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC100005D19"
                /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAAG 355
```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Labroidae; Cichlidae; Haplochromis.
REFERENCE
1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
Orf sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
1..417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stage="varied"
/clone_lib="HCEST library"
ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAATAACAATAG 60
Db 129 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAATAACAATAG 70
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 69 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 29
RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10 library Sterkiella
haplochromis genomic clone UUGC10006J13 R, genomic survey
sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0006 row: j column: 13
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES
Location/Qualifiers
1..491
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10006J13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAATAACAATAG 60
Db 412 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAATAACAATAG 353
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 352 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 312
RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES
Location/Qualifiers
1..495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

```

```
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 60
|
Db 62 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 121
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|
Db 122 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 162
|

RESULT 10
CC818374/c
LOCUS
DEFINITION
CC818374 495 bp DNA linear GSS 17-JUL-2003
100004807R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004807 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
AUTHORS
1 (bases 1 to 495)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
FEATURES
Location/Qualifiers
1. 495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004807"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 60
|
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 332
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|
```

```
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 60
|
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 333
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|
Db 332 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 292
|

RESULT 11
CC818523/c
LOCUS
DEFINITION
CC818523 496 bp DNA linear GSS 17-JUL-2003
100004L13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
AUTHORS
1 (bases 1 to 496)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
FEATURES
Location/Qualifiers
1. 496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 496;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 60
|
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAG 332
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|
```

```
Db 331 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 291
|||||
CC819854 503 bp DNA linear GSS 17-JUL-2003
100006N08R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100006N08 R, genomic survey
sequence.
RESULT 12
LOCUS CC819854/c
DEFINITION
ACCESSION
VERSION CC819854.1 GI:32900533
KEYWORDS
SOURCE
ORGANISM Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
AUTHORS 1 (bases 1 to 503)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL
COMMENT Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: N column: 08
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 503.
FEATURES
source
Location/Qualifiers
1..503
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100006N08"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 60
|||||
Db 410 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 351
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 350 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 310
|||||
Query Match 100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 60
|||||
Db 410 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 351
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 350 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 310
|||||
Query Match 100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 60
|||||
Db 412 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAATAAACAATAG 353
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 352 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 312
|||||
RESULT 14
LOCUS CC817128/c
DEFINITION CC817128 518 bp DNA linear GSS 17-JUL-2003
100002D21R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100002D21 R, genomic survey
sequence.
ACCESSION
VERSION CC817128.1 GI:32896415
KEYWORDS
SOURCE
ORGANISM Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
```

```

REFERENCE
AUTHORS      Stichotrichida; Oxytrichidae; Sterkiella.
TITLE        1 (bases 1 to 518)
JOURNAL      Dunn,D., Doak,T., Herrick,G. and Weiss,R.
COMMENT      Paired end reads from plasmid inserts of Oxytricha trifallax
              macronuclear chromosomes
              Unpublished (2003)
              Contact: Robert B. Weiss
              University of Utah Genome Center
              University of Utah
              Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
              84112, USA
              Tel: 801 585 5606
              Fax: 801 585 7177
              Email: ddunn@genetics.utah.edu
              Plate: 0002 row: D column: 21
              Seq primer: CACACAGGAAACAGCTATGACC
              Class: plasmid ends
              High quality sequence stop: 518.

FEATURES
source
1..518
    /organism="Sterkiella histriomuscorum"
    /mol_type="genomic DNA"
    /db_xref="taxon:94289"
    /clone="UUGC100002D21"
    /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
    /note="Vector: PWD42nv; Purified macronuclear chromosomal
    DNA from Oxytricha trifallax was blunt end-repaired with
    T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
    oligonucleotides were ligated to the blunt ends in high
    molar excess. Vector DNA was prepared from a derivative of
    PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
    derivative of plasmid R1. The vector was ligated with
    adaptors complementary to the insert adaptors and
    purified. The sheared, adapted mouse DNA was annealed to
    adapted vector DNA, and transformed into
    chemically-competent E. Coli XL10-Gold (Stratagene) cells
    and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 410 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 351

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 350 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 310

RESULT 15
LOCUS      CC817162/c
DEFINITION 519 bp DNA linear GSS 17-JUL-2003
            100002J19R Oxytricha plasmid UUGC100002J19 R, genomic survey
            histriomuscorum genomic clone UUGC100002J19 R, genomic survey
            sequence.
ACCESSION  CC817162
VERSION    CC817162
KEYWORDS   GSS.
SOURCE     Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM  Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;
            Stichotrichida; Oxytrichidae; Sterkiella.
            1 (bases 1 to 519)
            Dunn,D., Doak,T., Herrick,G. and Weiss,R.
            Paired end reads from plasmid inserts of Oxytricha trifallax
            macronuclear chromosomes
            Unpublished (2003)
            Contact: Robert B. Weiss
            University of Utah Genome Center

```

```

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0002 row: J column: 19
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 519.

FEATURES
source
1..519
    /organism="Sterkiella histriomuscorum"
    /mol_type="genomic DNA"
    /db_xref="taxon:94289"
    /clone="UUGC100002J19"
    /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
    /note="Vector: PWD42nv; Purified macronuclear chromosomal
    DNA from Oxytricha trifallax was blunt end-repaired with
    T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
    oligonucleotides were ligated to the blunt ends in high
    molar excess. Vector DNA was prepared from a derivative of
    PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
    derivative of plasmid R1. The vector was ligated with
    adaptors complementary to the insert adaptors and
    purified. The sheared, adapted mouse DNA was annealed to
    adapted vector DNA, and transformed into
    chemically-competent E. Coli XL10-Gold (Stratagene) cells
    and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 416 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 357

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 356 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:22:58
Job time : 961.667 secs

```

**THIS PAGE BLANK (USPTO)**



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	100	100.0	3853	6	AR098190	Sequence
2	100	100.0	3853	6	AR207832	Sequence
3	100	100.0	3853	6	BD009729	Sequence
4	100	100.0	3986	12	PCDNA32EO	Tissue sp
5	100	100.0	4026	6	AR098191	X90639 Cloning vec
6	100	100.0	4026	6	BD007833	Sequence
7	100	100.0	4026	6	BD009730	Sequence
8	100	100.0	4249	6	AR098192	Tissue sp
9	100	100.0	4249	6	AR207834	Sequence
10	100	100.0	4249	6	BD009731	Sequence
11	100	100.0	4341	6	AX8214	Tissue sp
12	100	100.0	4341	6	AX286570	Sequence 58
13	100	100.0	4597	6	AX060344	Sequence
14	100	100.0	4840	6	AX133940	Sequence
15	100	100.0	5053	6	BD238492	AX060344 Sequence
16	100	100.0	5070	6	BD234391	AX133940 Sequence
17	100	100.0	5082	6	AX91754	BD238492 Expressio
18	100	100.0	5082	6	BD085110	AX234391 Sequence
19	100	100.0	5082	6	BD085110	Sequence 10
20	100	100.0	5162	6	AX951626	Vertebra

**SOURCE**  
**ORGANIS**

```

Unclassified.
1 (bases 1 to 3853)
Antelman,D., Gregory,R.J. and Wills,K.N.
Retinoblastoma fusion proteins
Patent: US 6379927-A 5 30-APR-2002;
Location/Qualifiers
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 3
BD009729 3853 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified

REFERENCE
1 (bases 1 to 3853)
Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANUI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 209..862.
Location/Qualifiers
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995
LOCUS

REFERENCE
1 (bases 1 to 3853)
Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
Location/Qualifiers
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

PCDNA3ZEO 4026 bp DNA linear PAT 14-FEB-2001
LOCUS

DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:949972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
1
AUTHORS Peters,H., Hundhausen,T., Kroenke,M. and Marget,M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters,H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelstr.5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
source
1..3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pCDNA3ZEO"

misc_feature 1..2125
/notes="cloning vector (pCDNA3) (Invitrogen)"
889..994
/notes="multiple cloning site (MCS)"
2126..2796
/notes="cloning vector (PZeoSV) (Invitrogen)"
2797..3986
/notes="cloning vector (pCDNA3)"

ORIGIN
Query Match 100.0%; Score 100; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 5
AR098191 4026 bp DNA linear PAT 14-FEB-2001
LOCUS
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 4026)
Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
Location/Qualifiers
1..4026
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 6
AR207833
LOCUS 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES Location/Qualifiers
source 1. 4026
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

RESULT 7
BD009730
LOCUS 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
PT DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
CO H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. 4026
/organism='Unidentified'.
/organism="unassigned DNA"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES source
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 8
AR098192
LOCUS 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES Location/Qualifiers
source 1. 4249
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 9
AR207834
LOCUS 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
source 1. 4249
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
```



```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||

RESULT 14
AX133940
LOCUS          AX133940          4840 bp          DNA          linear          PAT 15-MAY-2001
DEFINITION    Sequence 1 from Patent WO0119853.
ACCESSION     AX133940
VERSION       AX133940.1 GI:14139881
KEYWORDS
SOURCE
ORGANISM      synthetic construct
               synthetic construct
               other sequences; artificial sequences.
REFERENCE
AUTHORS       Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE         Cell transfection
JOURNAL       THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES
source        1..4840
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="This sequence is artificial and is based on well
               established commercially available vectors that are cited
               with their vendor within the patent application"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||

RESULT 15
BD238492
LOCUS          BD238492          5053 bp          DNA          linear          PAT 17-JUL-2003
DEFINITION    Expression vectors for stimulating an immune response and methods
               of using the same.
ACCESSION     BD238492
VERSION       BD238492.1 GI:33048262
KEYWORDS      JP 2002520000-A/18.
SOURCE        synthetic construct
               synthetic construct
               other sequences; artificial sequences.
REFERENCE
AUTHORS       Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
               and Chesnut,R.W.
TITLE         Expression vectors for stimulating an immune response and methods
               of using the same
JOURNAL       Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904,15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
FT /organism="Artificial Sequence".
FEATURES
source      Location/Qualifiers
            1..5053
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
    |||||||

Search completed: July 14, 2005, 14:03:30
Job time : 749.127 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 gacggatcgggagatctccc.....ctgctccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	1506	12 ADM41035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADM41037	Adm41037 Cytochrome
4	100	100.0	2241	12 ADM41034	Adm41034 Human nuc
5	100	100.0	2294	12 ADM41036	Adm41036 Cytochrome
6	100	100.0	3853	2 AAV40006	Aav40006 Plasmid p
7	100	100.0	4026	2 AAV40007	Aav40007 Plasmid p
8	100	100.0	4249	2 AAV63466	Aav63466 Plasmid p
9	100	100.0	4341	2 AAG62391	Aag62391 Vector pV
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pV
11	100	100.0	4341	6 ABN83143	Abn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pEP2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA sequ
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rD

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 ADE21866	Ade21866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AAZ89476	Aaz89476 Transgeni
33	100	100.0	5446	6 AAS18619	Aas18619 Renilla l
34	100	100.0	5446	6 ABL53540	AbL53540 Vector pc
35	100	100.0	5446	12 ADN36314	Adn36314 Plasmid p
36	100	100.0	5458	6 ABL58494	AbL58494 Recombina
37	100	100.0	5458	6 ABL58493	AbL58493 Recombina
38	100	100.0	5543	6 ABK88868	Abk88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ade83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	AbL58489 Recombina
42	100	100.0	5614	6 ABL58490	AbL58490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AAI66195	Aai66195 Human FSH
45	100	100.0	5651	6 ABK40237	Abk40237 DNA encod

## ALIGNMENTS

## RESULT 1

ADMA41035

ID ADM41035 standard; DNA; 1506 BP.

AC ADM41035;

DT 17-JUN-2004 (first entry)

DE Fungus nucleotide sequence SEQ ID NO:3.

XX engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX tissue transplantation; human disease study; fungus; gene; ds.

OS Unidentified.

PN WO2004027029-A2.

PD 01-APR-2004.

PF 17-SEP-2003; 2003WO-US029251.

PR 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschoner WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a fetal non-human mammal host.

PS Disclosure; SEQ ID NO 3; 48pp; English.

XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the foetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a  
 CC mammalian host, and for producing chimeric mammals that can be used to  
 CC develop new drugs and vaccine, factors, drugs and tissues for  
 CC transplantation, also useful to study human diseases. The present  
 CC sequence represents a nucleotide sequence given in the Sequence Listing  
 CC of the present invention but not mentioned further within the  
 CC specification.

XX  
 SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;  
 Best Local Similarity 100.0%; Pred. No. 4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 2  
 ADH11349  
 ID ADH11349 standard; DNA; 1600 BP.

XX  
 AC ADH11349;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX  
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
 KW cell shape regulator; cell motility regulator; cell migration;  
 KW cell behaviour regulator; phenotype; signal transduction pathway;  
 KW signal transducing protein; signal integrator protein;  
 KW neuronal regeneration; revascularisation; wound healing;  
 KW chronic neurodegenerative disease; acute traumatic injury;  
 KW fibrotic disease; gene; ds.

XX  
 OS Unidentified.  
 XX  
 FN WO9824810-A2.  
 XX  
 PD 11-JUN-1998.  
 XX  
 PF 03-DEC-1997; 97WO-EP006956.  
 XX  
 PR 04-DEC-1996; 96GB-00025283.  
 XX

PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
 PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
 PI Geysen J, Bogaert TAOE;  
 XX  
 DR WPI; 1998-362411/31.  
 DR P-PSDB; ADH11350.  
 XX

PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.  
 PT promoting neuronal regeneration, treating chronic neuro-degenerative  
 PT diseases or acute traumatic injuries.  
 XX  
 PS Disclosure; Page 410-411; 479pp; English.

XX  
 CC The present invention describes a vertebrate protein homologue of an UNC-  
 CC 53 protein or Caenorhabditis elegans or a functional equivalent,  
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
 CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of  
 CC cell shape, motility, or the direction of cell migration for use as a  
 CC therapeutic; (7) a method for determination of whether a protein is an  
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
 CC motility or the direction of migration by contacting a host cell  
 CC expressing a homologue of UNC-53 and determining a change of phenotype;  
 CC (8) a method for identification of vertebrate homologues of C. elegans  
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
 CC a DNA library; and (9) a method for identification of a protein which is  
 CC active in the signal transduction pathway of a cell of which a vertebrate  
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
 CC antibody/homologue complex; and (iii) analysing such a complex to  
 CC identify any non-antibody protein bound to the complex. UNC-53 is a  
 CC signal transducing or signal integrator protein involved in controlling  
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate  
 CC homologues of UNC-53 can be used to promote neuronal regeneration,  
 CC revascularisation or wound healing, to treat chronic neurodegenerative  
 CC diseases or acute traumatic injuries or fibrotic diseases. The present  
 CC sequence is used in the exemplification of the present invention.

XX  
 SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 3  
 ADH41037  
 ID ADH41037 standard; DNA; 1782 BP.

XX  
 AC ADH41037;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX

DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.

XX  
 KW engrafting foreign replacement cell; implanting foreign replacement cell;  
 KW growth; differentiation; drug development; vaccine development;  
 KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX  
 OS Cytomegalovirus.  
 XX  
 FN WO2004027029-A2.  
 XX  
 PD 01-APR-2004.  
 XX  
 PF 17-SEP-2003; 2003WO-US029251.  
 XX  
 PR 19-SEP-2002; 2002US-0411790P.  
 XX

PA (XIME-) XIMEREX INC.  
 XX  
 PI Beschoner WE, Sosa CE, Thompson SC;  
 XX  
 DR WPI; 2004-295402/27.

XX  
 PT Engrafting foreign replacement cells within a fetal non-human mammal,  
 PT useful in producing chimeric mammals, comprises selectively destroying  
 PT native cells in a tissue of a fetal non-human mammal host.

XX  
 PS Disclosure; SEQ ID NO 5; 48pp; English.

XX  
 CC The present invention describes a method for engrafting foreign



CC replacement cells within a foetal non-human mammal, which comprises  
 CC selectively destroying native cells in a tissue of a foetal non-human  
 CC mammal host, where the number of maternal cells of the same tissue is not  
 CC substantially reduced, and implanting foreign replacement cells in the  
 CC tissue of the foetal non-human mammal host, where the foreign replacement  
 CC cells replace destroyed cells of the tissue. The method is useful for  
 CC facilitating growth and differentiation of foreign cells within a  
 CC mammalian host, and for producing chimeric mammals that can be used to  
 CC develop new drugs and vaccine, factors, drugs and tissues for  
 CC transplantation, also useful to study human diseases. The present  
 CC sequence represents a nucleotide sequence given in the Sequence Listing  
 CC of the present invention but not mentioned further within the  
 CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 100; DB 12; Length 1782;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTTGTGTT 100

RESULT 4  
 ADM41034  
 ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

DT 17-JUN-2004 (first entry)

XX Human nucleotide sequence SEQ ID NO:2.

XX engrafting foreign replacement cell; implanting foreign replacement cell;  
 XX growth; differentiation; drug development; vaccine development;  
 XX tissue transplantation; human disease study; human; gene; ds.

XX Homo sapiens.

XX WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,  
 XX useful in producing chimeric mammals, comprises selectively destroying  
 XX native cells in a tissue of a foetal non-human mammal host.

XX Disclosure; SEQ ID NO 2; 48pp; English.

XX The present invention describes a method for engrafting foreign  
 CC replacement cells within a foetal non-human mammal, which comprises  
 CC selectively destroying native cells in a tissue of a foetal non-human  
 CC mammal host, where the number of maternal cells of the same tissue is not  
 CC substantially reduced, and implanting foreign replacement cells in the  
 CC tissue of the foetal non-human mammal host, where the foreign replacement  
 CC cells replace destroyed cells of the tissue. The method is useful for  
 CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to  
 CC develop new drugs and vaccine, factors, drugs and tissues for  
 CC transplantation, also useful to study human diseases. The present  
 CC sequence represents a nucleotide sequence given in the Sequence Listing  
 CC of the present invention but not mentioned further within the  
 CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTTGTGTT 100

RESULT 5  
 ADM41036  
 ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

DT 17-JUN-2004 (first entry)

XX Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX engrafting foreign replacement cell; implanting foreign replacement cell;  
 XX growth; differentiation; drug development; vaccine development;  
 XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX Cytomegalovirus.

XX WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,  
 XX useful in producing chimeric mammals, comprises selectively destroying  
 XX native cells in a tissue of a foetal non-human mammal host.

XX Disclosure; SEQ ID NO 4; 48pp; English.

XX The present invention describes a method for engrafting foreign  
 CC replacement cells within a foetal non-human mammal, which comprises  
 CC selectively destroying native cells in a tissue of a foetal non-human  
 CC mammal host, where the number of maternal cells of the same tissue is not  
 CC substantially reduced, and implanting foreign replacement cells in the  
 CC tissue of the foetal non-human mammal host, where the foreign replacement  
 CC cells replace destroyed cells of the tissue. The method is useful for  
 CC facilitating growth and differentiation of foreign cells within a  
 CC mammalian host, and for producing chimeric mammals that can be used to  
 CC develop new drugs and vaccine, factors, drugs and tissues for  
 CC transplantation, also useful to study human diseases. The present  
 CC sequence represents a nucleotide sequence given in the Sequence Listing  
 CC of the present invention but not mentioned further within the  
 CC specification.



```
FT CDS complement(3032..3850)
FT /*tag= f
FT /*note= "AMP-ORF"
XX
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26..
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 6; 91pp; English.
XX
XX This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX vector. Plasmid pCTMI has been used as a vector for the expression of
XX fusion proteins of the invention that comprise retinoblastoma protein
XX (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX fusion proteins, particularly expressed from gene therapy vectors, are
XX used to treat hyperproliferative conditions, specifically cancer
XX (particularly of the bladder) or restenosis. They are more effective in
XX repressing transcription of the E2F promoter than RB alone and cause cell
XX -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX OS field.)
XX
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 100; DB 2; Length 4026;
XX Best Local Similarity 100.0%; Pred. No. 5.3e-26;
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
XX |||||
XX Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
XX |||||
XX Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTTT 100
XX |||||
XX Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTTT 100
XX |||||
XX
XX RESULT 8
XX AAV63466
XX ID AAV63466 standard; DNA; 4249 BP.
XX
XX AC AAV63466;
XX
XX 27-AUG-2003 (revised)
XX 15-FEB-1999 (first entry)
XX
XX DE Plasmid pCTMIE.
XX
XX E2F; transcription factor; human; retinoblastoma protein RB;
XX bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX thyroid hyperplasia; Grave's disease; psoriasis;
XX benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX
XX OS Human cytomegalovirus.
XX OS mastadenovirus.
XX OS unidentified bacteriophage; T7.
XX OS unidentified bacteriophage; SP6.
```

```
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
XX Key Location/Qualifiers
XX Promoter 209..864
XX /*tag= a
XX /*note= "CMV promoter"
XX misc_feature 907..1074
XX /*tag= b
XX /*function= "tripartite leader sequence"
XX intron 1081..1145
XX /*tag= c
XX /*note= "hybrid SV40 late intron"
XX mRNA 1164..1366
XX /*tag= d
XX /*note= "early mRNA"
XX enhancer 1261..1332
XX /*tag= e
XX /*note= "72 bp tandem repeat enhancer"
XX enhancer 1333..1404
XX /*tag= f
XX /*note= "72 bp tandem repeat enhancer"
XX misc_binding 1366
XX /*tag= g
XX /*note= "T antigen binding site"
XX intron 1372..1478
XX /*tag= h
XX /*note= "hybrid SV40 late intron"
XX promoter 1530..1545
XX /*tag= i
XX /*note= "SP6 promoter"
XX misc_feature 2075..4249
XX /*tag= j
XX /*note= "pUC19 backbone H3 to AatII"
XX CDS complement(3255..4113)
XX /*tag= k
XX /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 8; 91pp; English.
XX
XX This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX amplified product with BglII and inserting into BamHI-digested plasmid
XX pCTMI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX expression of fusion proteins of the invention that comprise
XX retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX (see AAW62464). Such fusion proteins, particularly expressed from gene
XX therapy vectors, are used to treat hyperproliferative conditions,
XX specifically cancer (particularly of the bladder) or restenosis. They are
XX more effective in repressing transcription of the E2F promoter than RB
XX alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX SQ Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;
```

Query Match 100.0%; Score 100; DB 2; Length 4249;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9  
AAQ62391  
ID AAQ62391 standard; DNA; 4341 BP.

XX AC AAQ62391;  
XX DT 25-MAR-2003 (revised)  
XX DT 18-NOV-1994 (first entry)  
XX DE Vector pVAC1.  
XX KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;  
XX KW fusion protein; pSfi/NotI-tag1; pElB leader; human; immunoglobulin; VHI;  
XX KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;  
XX KW vaccine; ss.  
XX OS Synthetic.

XX FH Key Location/Qualifiers  
XX FT misc\_RNA complement (1..775)  
XX FT /\*tag= c  
XX FT /note= "Claim 9"  
XX FT misc\_RNA 606..780  
XX FT /\*tag= b  
XX FT /note= "Claim 8"  
XX FT misc\_RNA 606..716  
XX FT /\*tag= a  
XX FT /note= "Claim 7"

XX W09408008-A1.  
XX 14-APR-1994.  
XX 04-OCT-1993; 93WO-GB002054.  
XX 02-OCT-1992; 92GB-00020808.  
XX (MEDI-) MEDICAL RES COUNCIL.  
XX PA Hawkins RE, Russell SJ, Stevenson FK, Winter GP;  
XX PI WPI; 1994-135575/16.  
XX DR  
XX Modulating immune response to a disease marker - by administering a  
XX vector which expresses the disease marker to interact with the immune  
XX system.  
XX PS Claim 10; Fig 7; 77pp; English.

XX This sequence represents the vector pVAC1. This vector is based on the  
XX commercially available vector pRC/RSV. Leader sequences and termination  
XX signals were introduced into the vector to allow for production of fusion  
XX proteins. The vector, pSfi/NotI-tag1, was modified to replace the pElB  
XX leader with the human immunoglobulin VHI leader sequence that permits the  
XX encoding of an sFII cloning site without modification of the amino acid  
XX sequence. This fragment was then cloned as an EcoRI/Bln-HindIII  
XX fragment into NotI/Bln-HindIII cut vector pRC/RSV to give pVAC1. The  
XX single chain Fv for an individual patient can be inserted within the VHI  
XX leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid  
CC vaccine and it induces a strong humoral response to the antibody moiety  
CC in BALB/c mice. (Updated on 25-MAR-2003 to correct FN field.)  
XX

SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 2; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 10  
AAS17704  
ID AAS17704 standard; DNA; 4341 BP.

XX AC AAS17704;  
XX DT 12-MAR-2002 (first entry)  
XX DE Vector pVAC1 encoding a DNA vaccine.  
XX KW Cytostatic; vaccine; tetanus toxin; Frc; tumour; CTL; PCR primer; pVAC1;  
XX KW ds.  
XX OS Clostridium tetani.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX OS Cauliflower mosaic virus.  
XX W0200179510-A1.  
XX 25-OCT-2001.  
XX 17-APR-2001; 2001WO-GB001719.  
XX 17-APR-2000; 2000GB-00009470.  
XX (CANC-) CANCER RES VENTURES LTD.  
XX Rice J, Stevenson F;  
XX WPI; 2002-086370/09.  
XX Nucleic acid construct, useful to immunize against various diseases  
XX including cancer, expresses the first domain of tetanus toxin Frc fused  
XX to a disease peptide antigen to provide a vaccine.  
XX Disclosure; Fig 4; 71pp; English.

XX The invention relates to a nucleic acid construct for delivery into  
XX living cells in vivo, to induce an immune response to a disease peptide  
XX antigen, where the construct directs expression of a fusion protein  
XX comprising the peptide antigen and the first domain of Frc. Also included  
XX are a nucleic acid vector comprising the above construct, a host cell  
XX comprising the above construct or vector and a method of producing a  
XX nucleic acid construct for inducing an immune response. The method  
XX comprises identifying a nucleic acid sequence encoding a disease peptide  
XX antigen comprising epitopes characteristic of the disease, cloning the  
XX nucleic acid sequence, introducing the cloned nucleic acid into a vector  
XX which allows the antigen to be expressed as a fusion with a first domain  
XX Frc from tetanus toxin, and optionally isolating the construct from the  
XX vector. The construct or vector is used as a vaccine to induce an immune  
XX response, particularly to tumour antigens. The present sequence is vector  
XX pVAC1 which encodes a vaccine of the invention

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 11  
ID ABN83143 standard; DNA; 4341 BP.  
XX  
AC ABN83143;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Plasmid pVAC1 complete sequence.  
XX  
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;  
KW cancer; B cell malignancy; ds.  
XX  
OS Synthetic.  
XX  
PN WO200240513-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-GB005142.  
XX  
PR 20-NOV-2000; 2000GB-00028319.  
XX  
PA (CANC-) CANCER RES VENTURES LTD.  
XX  
PI Savelyeva N, Stevenson F;  
XX  
DR WPI; 2002-500202/53.  
XX  
PS Nucleic acid construct for delivery into living cells as a vaccine,  
PT useful for treating e.g. cancer, directs the expression of a fusion  
PT protein comprising an antigen and an adjuvant sequence derived from a  
PT plant viral coat protein.  
XX  
XX  
XX Example 3; Fig 7; 84pp; English.  
XX  
CC The invention relates to a novel nucleic acid construct for inducing an  
CC immune response in vivo to an antigen, capable of directing the  
CC expression of a fusion protein that comprises an antigen and an adjuvant  
CC sequence derived from a plant viral coat protein. The construct of the  
CC invention has cytostatic and virucide activity. The nucleic acid  
CC construct is useful for inducing an immune response in a patient, for  
CC vaccinating a patient against an infectious disease caused by an antigen  
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a  
CC patient with a predisposition to cancer and for treating a patient having  
CC a B cell malignancy, where the construct is encapsulated, and optionally,  
CC a second nucleic acid sequence encoding a further immunomodulatory  
CC polypeptide is administered to the patient. The construct is also useful  
CC in medical treatment, and in the preparation of a vaccine for treating or  
CC preventing a disease state associated with the antigen. The sequence  
CC shows the complete sequence of vector pVAC1  
XX  
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 12  
ID AAF24901 standard; DNA; 4597 BP.  
XX  
AC AAF24901;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.  
XX  
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;  
KW myocardial ischemia; cardiac angiogenesis; haemophilia;  
KW vascular endothelial growth factor; VEGF; ss.  
XX  
OS Synthetic.  
XX  
PN WO200078358-A2.  
XX  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000WO-US016837.  
XX  
PR 18-JUN-1999; 99US-0140260P.  
XX  
PA (COLL-) COLLABORATIVE GROUP LTD.  
XX  
PI Chen W;  
XX  
DR WPI; 2001-071363/08.  
XX  
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial  
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic  
PT acids crosslinked to nucleic acids.  
XX  
XX Example 1; Page 36-38; 38pp; English.  
XX  
CC The specification describes a microsphere comprising dihydrazide  
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The  
CC microspheres cause reduced inflammatory responses, and have increased  
CC safety and biodegradability. The microspheres are useful for transfecting  
CC a cell of a subject and for treating a subject having myocardial  
CC ischemia, by increasing cardiac angiogenesis. They are also useful for  
CC treating haemophilia. The present sequence represents the plasmid  
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is  
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a  
CC vascular endothelial growth factor (VEGF)  
XX  
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 4; Length 4597;  
Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 13  
AAD39652



XX 27-FEB-2002; 2002US-0360274P.  
XX (MERI ) MERCK & CO INC.  
XX Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;  
XX Miller MD, Register B, Shi X, Simon AJ, Zuck PD;  
XX WPI; 2003-689968/65.  
XX DNA encoding a fusion protein of amyloid precursor protein, useful in  
XX screening for anti-Alzheimer agents, comprises a fused transcription  
XX factor.  
XX Disclosure; Fig 32B-F; 193pp; English.  
XX The present invention describes a DNA molecule (I) that encodes a fusion  
XX protein (FP) comprising: (i) an amino acid sequence of amyloid precursor  
XX protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a  
XX transcription factor (TF), fused in frame to the C-terminus of (i). Also  
XX described: (1) an expression vector containing (I); (2) a eukaryotic cell  
XX containing (I); and (3) methods for identifying a compound (A) that  
XX inhibits processing of APP, using the cells of (2). (I) has neurotropic and  
XX neuroprotective activities. (1) can be used to produce eukaryotic cells  
XX that express FP and are useful in screening for agents that inhibit  
XX processing of APP. The agents are potentially useful for the treatment or  
XX prevention of Alzheimer's disease. Cells that express FP can screen for  
XX inhibitors of: (a) beta- and gamma-secretases; and (b)  
XX cytoplasmic/extracellular APP signaling in a single assay. Cell-based  
XX assays may be free of interference from alpha-secretase activity and are  
XX homogeneous (no chromatography, immunoprecipitation or washing required)  
XX so well suited to high-throughput screening. The present sequence  
XX represents a plasmid nucleotide sequence from the present invention.  
SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 10; Length 5015;  
Best Local Similarity 100.0%; Pred. No. 5.6e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACGGATCGGGAGATCTCCGATCCCTATGGTCGACTTCAGTACATCTGCTCTGATG 60  
Db 1 GACGGATCGGGAGATCTCCGATCCCTATGGTCGACTTCAGTACATCTGCTCTGATG 60  
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCTGCTTGTTGTT 100  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCTGCTTGTTGTT 100

Search completed: July 14, 2005, 07:01:38  
Job time : 141.038 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)

3997.736 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_1\_100

Perfect score: 100

Sequence: 1 gacggatcgaggatctccc.....ctgctccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hcc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	60.0	602	8	B67169 CPG0047A Cp
2	55.6	55.6	694	8	B2052929 jnr13g03.
3	55.6	55.6	696	8	B2050328 jnr42c12.
4	55.6	55.6	717	8	B2054067 jnr38b09.
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK119397 212c09.p1
9	53.4	53.4	766	7	CK120360 207j04.p1
10	53.4	53.4	788	7	CK117844 209p08.p1
11	53.4	53.4	898	9	CL141237 ISB1-118J
12	53.4	53.4	899	9	CL140877 ISB1-118B
13	53.4	53.4	1009	9	CL123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbeb0049M
15	53	53.0	675	8	B2051815 jnr57d03.
16	53	53.0	679	8	B2052857 jnr13g03.
17	53	53.0	700	8	B2050646 jnr66f08.
18	53	53.0	701	8	B2052015 jnr56b03.
19	53	53.0	708	8	B2054793 jnr33g03.
20	53	53.0	709	8	B2053587 jnr98d01.
21	53	53.0	712	8	B2054005 jnr38b09.
22	52.8	52.8	451	8	AQ863966 nbeb0022E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	B2049113 jnr21d02.

25	52.4	52.4	708	8	BZ050047
26	51.6	51.6	328	9	CC819886
27	51.6	51.6	351	9	CC818492
28	51.6	51.6	358	9	CC817661
29	51.6	51.6	364	9	CC817805
30	51.6	51.6	364	9	CC818511
31	51.6	51.6	364	9	CC818574
32	51.6	51.6	364	9	CC819049
33	51.6	51.6	369	9	CC817069
34	51.6	51.6	374	9	CC817074
35	51.6	51.6	374	9	CC820036
36	51.6	51.6	395	9	CC817652
37	51.6	51.6	403	9	CC817682
38	51.6	51.6	403	9	CC817837
39	51.6	51.6	414	9	CC819240
40	51.6	51.6	419	9	CC818384
41	51.6	51.6	420	9	CC817834
42	51.6	51.6	426	9	CC817720
43	51.6	51.6	437	9	CC819820
44	51.6	51.6	441	9	CC818421
45	51.6	51.6	443	9	CC817769

#### ALIGNMENTS

RESULT 1  
B67169  
LOCUS B67169 602 bp DNA linear GSS 12-MAY-2000  
DEFINITION CPG0047A CplOWAgDNA2 Cryptosporidium parvum genomic, genomic survey sequence.  
ACCESSION B67169  
VERSION B67169.1 GI:2642750  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum  
ORGANISM Cryptosporidium parvum  
REFERENCE 1 (bases 1 to 602)  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
MEDLINE 20183851  
PUBMED 10717299  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malariad@itsa.ucsf.edu  
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.  
Seq primer: T7  
Class: Shotgun  
High quality sequence stop: 602.  
Location/Qualifiers  
1. .602  
/organism="Cryptosporidium parvum"  
/mol\_type="genomic DNA"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/lab\_host="E. coli XL2 Blue MRF"  
/clone\_lib="CplOWAgDNA2"  
/notes="Vector: pCR-Script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

# ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;

Best Local Similarity 100.0%; Pred. No. 2.4e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAAATCTGCTCTGATGCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

Db 1 CAGTACAAATCTGCTCTGATGCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTT 60

## RESULT 2

BZ052929/c

LOCUS

DEFINITION jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ052929

VERSION BZ052929.1 GI:23654922

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr13 row: 9 column: 03

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

## FEATURES

source

1..694

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 62

Db 324 CGGATCGATAGGTCCTCGGACTAGTTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 265

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

## RESULT 3

BZ050328

LOCUS

DEFINITION jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ050328

VERSION BZ050328.1 GI:23649718

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: C column: 12

Seq primer: -21UPpOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

## FEATURES

source

1..696

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 62

Db 45 CGGATCGATAGGTCCTCGGACTAGTTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 104

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

## RESULT 4

BZ054067/c

LOCUS

DEFINITION jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ054067

VERSION BZ054067.1 GI:23657216

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

TITLE Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
JOURNAL Whole genome shotgun reads from *Brassica oleracea*  
COMMENT Unpublished (2002)  
Contact: Richard K. Wilson

Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Plate: jnr38 row: b column: 09  
Seq primer: -28RPOT reverse  
Class: shotgun  
High quality sequence start: 87  
High quality sequence stop: 543.  
Location/Qualifiers

## FEATURES

source

1. .717  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea001"  
/note="Vector: pOTw13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using *Brassica oleracea* T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;  
Best Local Similarity 77.9%; Pred. No. 9.1e-09;  
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCC 62

Db 248 CGGATCGATAGGTCCCTCGACTAGTATTATGTCGACTCTCAGTACAACTGCTCTGATGCC 189

Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88

Db 188 GCATAGTTAAGCCAGCCCGACACC 163

## RESULT 5

AW409112

LOCUS

DEFINITION AW409112 348 bp mRNA linear EST 31-DEC-2000  
sal10h5 Salivary Gland Library Homo sapiens cDNA, mRNA sequence.

ACCESSION AW409112

VERSION AW409112.1 GI:11999687

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

AUTHORS

1 (bases 1 to 348)  
Bergheim, A., Ogilvie, E., Arndt, S., Napier, H., Taylor, M., Lovett, M.,  
Simmons, A., Hide, W. and Ramsay, M.

TITLE A high density transcript map between markers D8S550 and D8S1759 on  
8p22-23, using cDNA direct selection

JOURNAL Unpublished (2000)

## COMMENT

Contact: Ramsay M  
Department of Human Genetics  
South African Institute For Medical Research  
P.O.Box 1038, Johannesburg, Gauteng, 2000, South Africa  
Fax: 2711 489 9226  
Email: micheler@mail.saimr.wits.ac.za.

## FEATURES

source

1. .348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Salivary Gland"  
/clone\_lib="Salivary Gland Library"  
/note="Vector: pAMP10"

## ORIGIN

## Query Match

Best Local Similarity

Matches 62; Conservative

Qy 11 GAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTT 70

Db 65 GCGGTATACACACCGCATATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTT 124

Qy 71 AAGCAGTATCTGCTCC 87

Db 125 AAGCAGTATACACTCC 141

## RESULT 6

AL715724/c

LOCUS

DEFINITION AL715724 343 bp mRNA linear EST 18-APR-2002  
rerio cDNA clone BNOAA018ZF12 5', mRNA sequence.

ACCESSION AL715724

VERSION AL715724.1 GI:20180327

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 343)

AUTHORS Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,

Hardelin, J.P., Weissenbach, J. and Petit, C.

TITLE A subtracted cDNA library from the zebrafish (Danio rerio)

embryonic inner ear

unpublished (2002)

JOURNAL Contact: Genoscope

COMMENT Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES source

1. .343

/organism="Danio rerio"

/mol\_type="mRNA"

/db\_xref="taxon:7955"

/clone\_lib="BNOAA018ZF12"

/tissue\_type="inner ear"

/dev\_stage="embryonic"

/clone\_lib="Danio rerio embryonic inner ear subtracted

cDNA"

/note="subtracted cDNA library"

## ORIGIN

## Query Match

Best Local Similarity

Matches 60; Conservative

Qy 17 TCCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76

Db 280 TTACACCGATATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 221

Qy 77 GTATCTGCTCC 87

Db 220 GTATACACTCC 210

## RESULT 7

AL714571/c

LOCUS

DEFINITION AL714571 345 bp mRNA linear EST 18-APR-2002  
rerio cDNA clone BNOAA007ZC02 5', mRNA sequence.

ACCESSION AL714571

VERSION AL714571.1 GI:20179174

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 345) Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin,J.P., Weissenbach,J. and Petit,C. A subtracted cDNA library from the zebrafish (Danio rerio) embryonic inner ear Unpublished (2002) Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: segref@genoscope.cns.fr, web : www.genoscope.cns.fr.	Location/Qualifiers 1. .345 /organism="Danio rerio" /mol_type="mRNA" /db_xref="taxon:7955" /clone="BN0AA007ZC02" /tissue_type="inner ear" /dev_stage="embryonic" /clone_lib="Danio rerio embryonic inner ear subtracted cDNA" /note="subtracted cDNA library"	Query Match 53.4%; Score 53.4; DB 1; Length 345; Best Local Similarity 84.5%; Pred. No. 4.8e-08; Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;	QY 17 TCCGATGCCCTATGGTCGACTCTCAGTACAATCTCTGATGCCGATGTTAAGCCA 76 Db 280 TTCACCGCATATGGTCGACTCTCAGTACAATCTCTGATGCCGATGTTAAGCCA 221 QY 77 GTATCTGCTCC 87 Db 220 GTATACACTCC 210	RESULT 8 CKI19397/c LOCUS DEFINITION 5-PRIME, mRNA sequence. ACCESSION VERSION KEYWORDS SOURCE ORGANISM Arabidopsis thaliana (thale cress) Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 761) Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B. Generation of a cDNA expression library from Arabidopsis inflorescence meristem Unpublished (2003) Contact: Birgit Kersten Plant Protein Chip Group, Department Lehrach Max-Planck-Institute for Molecular Genetics Innstr. 73, D-14195 Berlin, Germany Tel: +49(0)30/84131648 Fax: +49(0)30/84131128 Email: Kersten@molgen.mpg.de Insert Length: 761 Std Error: 0.00 Plate: 212 row: 0 column: 9 Seq primer: pQE65. Location/Qualifiers 1. .761 /organism="Arabidopsis thaliana" /mol_type="mRNA" /ecotype="Columbia" /db_xref="GABI:954234" /db_xref="taxon:3702"	761 bp mRNA linear EST 01-JUN-2004 CKI19397 212009.pl ALM1 Arabidopsis thaliana cDNA clone MPMGp2011009212 5-PRIME, mRNA sequence. CKI19397 CKI19397.1 GI:47829713 EST. Arabidopsis thaliana (thale cress) Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 761) Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B. Generation of a cDNA expression library from Arabidopsis inflorescence meristem Unpublished (2003) Contact: Birgit Kersten Plant Protein Chip Group, Department Lehrach Max-Planck-Institute for Molecular Genetics Innstr. 73, D-14195 Berlin, Germany Tel: +49(0)30/84131648 Fax: +49(0)30/84131128 Email: Kersten@molgen.mpg.de Insert Length: 761 Std Error: 0.00 Plate: 212 row: 0 column: 9 Seq primer: pQE65. Location/Qualifiers 1. .761 /organism="Arabidopsis thaliana" /mol_type="mRNA" /ecotype="Columbia" /db_xref="GABI:954234" /db_xref="taxon:3702"
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Site\_2: NotI; About 1 week after bolting, cDNA synthesis using Superscript<sup>TM</sup>-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearrayed into the sublibrary (plate numbers begin with 201) containing 5,000 putative expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 766;  
Best Local Similarity 84.5%; Pred. No. 5.6e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTGATGCTCAGTACATCTGCTCTGATGCCGATAGTTAAGCCA 76  
Db 679 TTCACCGCATATGGTGCACTCTCAGTACATCTGCTCTGATGCCGATAGTTAAGCCA 620  
Qy 77 GTATCTGCTCC 87  
Db 619 GTATACACTCC 609

## RESULT 10

CK117844/c  
LOCUS CK117844 788 bp mRNA linear EST 01-JUN-2004  
DEFINITION 209p08.p1 AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011P08209  
5-PRIME, mRNA sequence.

ACCESSION CK117844

VERSION CK117844.1 GI:47828160

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 788)

REFERENCE Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

CONTACT Birgit Kersten

Plant Protein Chip Group, Department Lehrach

Max-Planck-Institute for Molecular Genetics

Instr. 73, D-14195 Berlin, Germany

Tel: +49 (0)30/84131648

Fax: +49 (0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 788 Std Error: 0.00

Plate: 209 row: P column: 8

Seq primer: pQE65.

Location/Qualifiers

1..788

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Columbia"

/db\_xref="GABI:953578"

/db\_xref="taxon:3702"

/clone="MPMPGP2011P08209"

/tissue\_type="inflorescence meristem"

/dev\_stage="about one week after bolting"

/lab\_host="E. coli SCS-1/pSE111"

/clone\_lib="AtM1"

/note="Vector: pQE-3ONAST-attB (AY386205); Site 1: SalI;

Site 2: NotI; About 1 week after bolting, cDNA synthesis

using Superscript<sup>TM</sup>-system (Invitrogen) with an

oligo(dT)-primer containing NotI restriction site and a

SalI adapter. The main library (plate numbers begin with

1) of 38,000 clones was rearrayed into the sublibrary

(plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 788;  
Best Local Similarity 84.5%; Pred. No. 5.7e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTGATGCTCAGTACATCTGCTCTGATGCCGATAGTTAAGCCA 76  
Db 514 TTCACCGCATATGGTGCACTCTCAGTACATCTGCTCTGATGCCGATAGTTAAGCCA 455  
Qy 77 GTATCTGCTCC 87  
Db 454 GTATACACTCC 444

## RESULT 11

CL141237/c

LOCUS CL141237 898 bp DNA linear GSS 05-JAN-2004

DEFINITION ISB1-118117\_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118117, genomic survey sequence.

ACCESSION CL141237

VERSION CL141237.1 GI:40634872

KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; xenopodinae; Xenopus; Silurana.  
1 (bases 1 to 898)

REFERENCE Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

JOURNAL Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert Length: 75000 Std Error: 0.00

Seq primer: T7 TAATACGACTCACTATAGG

Class: BAC ends

High quality sequence start: 4

High quality sequence stop: 742.

Location/Qualifiers

1..898

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/db\_xref="taxon:8364"

/clone="ISB1-118117"

/clone\_lib="ISB1"

/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

Library Segment 1"

```

LOCUS      CL140877                      899 bp    DNA        linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
            genomic survey sequence.
ACCESSION   CL140877
VERSION     CL140877.1 GI:40634512
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 899)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E., and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   source
            Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 899;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY      17 TCCCGATCCCTATGCTGCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
        |||||
        195 TTCACCGCATATGTCACCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 136
QY      77 GTATCTGCTCC 87
        |||||
        135 GTATACACTCC 125
        |||||

RESULT 13
LOCUS      CL123953/c
DEFINITION ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
            genomic survey sequence.
ACCESSION   CL123953
VERSION     CL123953.1 GI:40617588
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1009)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E., and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nbeb0049M21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /note="Vector: pBACindigo; Site_1: EcorI; Site_2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   source
            Location/Qualifiers
                1..1009
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-84J15"
                /clone_lib="ISB1"
                /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY      17 TCCCGATCCCTATGCTGCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
        |||||
        252 TTCACCGCATATGTCACCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 193
QY      77 GTATCTGCTCC 87
        |||||
        192 GTATACACTCC 182
        |||||

RESULT 14
LOCUS      AQ914559                      814 bp    DNA        linear    GSS 02-DEC-1999
DEFINITION nbeb0049M21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
            cultivar-group) genomic clone nbeb0049M21r, genomic survey
            sequence.
ACCESSION   AQ914559
VERSION     AQ914559.1 GI:6511075
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 814)
AUTHORS    Wing, R.A. and Dean, R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL    Unpublished (1998)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nbeb0049M21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /note="Vector: pBACindigo; Site_1: EcorI; Site_2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, the Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center ([www.genome.clemson.edu](http://www.genome.clemson.edu))."

## ORIGIN

Query Match 53.2%; Score 53.2; DB 8; Length 814;  
 Best Local Similarity 78.0%; Pred. No. 6.7e-08;  
 Matches 64; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 7 TGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATGCCGAT 66  
 |||||  
 Db 279 TGGGGGATTTTCACACCGCATATGGTGCATCTCAGTACATCTGCTCTGATGCCGAT 338  
 |||||

QY 67 AGTTAAGCCAGTATCTGCTCC 88  
 |||||

Db 339 AGTTAAGCCAGCCGACACCC 360  
 |||||

## RESULT 15

BZ051815  
 LOCUS  
 DEFINITION jnr57d03.b1 B.oleracea001 Brassica oleracea genomic, genomic survey sequence.  
 ACCESSION BZ051815  
 VERSION BZ051815.1 GI:23652690  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 675)  
 Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Naeh,W., Rabinowicz,P.D. and Wilson,R.K.  
 Whole genome shotgun reads from Brassica oleracea  
 Unpublished (2002)  
 Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 Plate: jnr57 row: d column: 03  
 Seq primer: -21UPPOT forward  
 Class: shotgun  
 High quality sequence start: 29  
 High quality sequence stop: 94.  
 Location/Qualifiers

## FEATURES

source  
 1..675  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOFw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T0100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

## ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;  
 Best Local Similarity 75.6%; Pred. No. 7.6e-08;

Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CCGATCGGGAGATCTCCGATCCCTATGGTGGACTCTCAGTACATCTGCTCTGATGCC 62  
 |||||

Db 53 CGGNACGATAGGTCCCTGGACTAGTTATGGTGGACTCTCAGTACATCTGCTCTGATGCC 112  
 |||||

QY 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||

Db 113 GCATAGTTAAGCCAGCCGACACCC 138  
 |||||

Search completed: July 14, 2005, 23:22:58  
 Job time : 952.146 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_3930\_4030

Perfect score: 101

Sequence: 1 ctcgagtctagagggccgt.....tccccgtgcttcttgac 101

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	280	6	CQ788637
2	101	100.0	4291	12	AF302189
3	101	100.0	5082	6	A91754
4	101	100.0	5082	6	BD085110
5	101	100.0	5432	6	BD234590
6	101	100.0	5432	6	AX026821
7	101	100.0	5650	6	AX226281
8	101	100.0	5731	6	AX202478
9	101	100.0	5771	12	AF060226
10	101	100.0	6050	6	AX195205
11	101	100.0	6082	6	AR278592
12	101	100.0	6082	6	AR367288
13	101	100.0	6082	6	AR400324
14	101	100.0	6082	6	AR405591
15	101	100.0	6082	6	AR563971
16	101	100.0	6082	6	AX141045
17	101	100.0	6082	6	AX200905
18	101	100.0	6082	6	AX267561
19	101	100.0	6253	6	AR031374

20	101	100.0	6253	6	BD009742
21	101	100.0	6338	6	BD134374
22	101	100.0	6338	6	AR428934
23	101	100.0	6365	6	AX513181
24	101	100.0	6567	6	AX128350
25	101	100.0	6623	6	AX128345
26	101	100.0	6639	6	AX128351
27	101	100.0	6649	6	AX180726
28	101	100.0	6695	6	AX128347
29	101	100.0	6695	6	AX128353
30	101	100.0	6695	6	AX128354
31	101	100.0	6746	6	AX128344
32	101	100.0	6759	6	CQ788635
33	101	100.0	6759	6	CQ788642
34	101	100.0	6801	6	AX128352
35	101	100.0	6801	6	AX128355
36	101	100.0	6818	6	AX128346
37	101	100.0	6828	6	AX128340
38	101	100.0	6833	6	AX128349
39	101	100.0	6900	6	AX128341
40	101	100.0	6956	6	AX128348
41	101	100.0	7038	6	AX128342
42	101	100.0	7108	6	E36262
43	101	100.0	7108	6	AX001326
44	101	100.0	7231	6	BD268252
45	101	100.0	7427	6	CQ768745

#### ALIGNMENTS

RESULT 1  
LOCUS CQ788637 280 bp DNA linear PAT 24-MAR-2004  
DEFINITION Sequence 117 from Patent WO2004020643.  
ACCESSION CQ788637  
VERSION CQ788637.1 GI:45723394  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Curtiss,R.I. and Kong,W.  
TITLE Regulated bacterial lysis for gene vaccine vector delivery and antigen release  
JOURNAL Patent: WO 2004020643-A 117 11-MAR-2004;  
WASHINGTON UNIVERSITY (US)  
FEATURES  
Location/Qualifiers  
source 1..280  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/notes="Description of Artificial Sequence: Multiple cloning site of pYA3650"

#### ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 280;  
Best Local Similarity 100.0%; Pred.No. 2.6e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTCG 60  
Dy 126 CTCGAGTCTAGAGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTCG 185  
Qy 61 CAGCCATCTGTTTGGCCCTCCCGTCCTTCCCTTGAC 101  
Dy 186 CAGCCATCTGTTTGGCCCTCCCGTCCTTCCCTTGAC 226

#### RESULT 2

AF302189 4291 bp DNA linear SYN 11-DEC-2001  
LOCUS AF302189  
DEFINITION Synthetic construct UOATP2 fusion protein gene, complete cds.

```

ACCESSION AF302189
VERSION AF302189.1 GI:17483748
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 4291)
AUTHORS Zullo,S.J., Parks,W.T., Chloupkova,M., Penton,W.A., Merrill,C.R. and Eisenstadt,J.M.
TITLE Expression of oligomycin resistance (olir) in CHO cells following transposition of the mitochondrial DNA-encoded olir ATPase 6 gene to the nuclear genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4291)
AUTHORS Zullo,S.J., Eisenstadt,J.M., Penton,W.A. and Merrill,C.R.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2000) Lab. Biochemical Genetics, NIMH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
    source
        1..4291
            Location/Qualifiers
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                /plasmid="pUOATP2"
        483..1259
            CDS
                /note="contains Home sapiens ornithine transcarbamylase leader sequence and Crictetus griseus mtATPase6 gene"
                /codon_start=1
                /transl_table=11
                /product="UOATP2 fusion protein"
                /protein_id="AAL40188.1"
                /db_xref="GI:17483749"
                /translation="MLNLRLLNNAAFRNHNPVNRFCGQPLQNNENLFSFIFPP
                TLMGLPIILIMPFPVMTSSKRLVNNRFTFOWLKILITKQMAIHSPKGTWSL
                MLASLIIFIGTINLGLLPHFTPTTOLSMNLGNAIPWAGAVILGFRHKWDSLAHF
                LPQGTPIPLPMLVIKITSLSFIQPMALAVRLTANITAGHLMLHLLGGATVLTLSISL
                PTAMITFILLMLILEFAVALIQAYVFTLLVSLYLDNT"
ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 4291;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 60
Db 1299 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 1358

Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 101
Db 1359 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 1399

RESULT 3
A91754
LOCUS A91754 5082 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent WO9824810.
ACCESSION A91754
VERSION A91754.1 GI:6740671
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS Bogaert,T.A. and Deraeymaeker,M.
TITLE VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS
JOURNAL Patent: WO 9824810-A 10 11-JUN-1998;
BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEMYAEKER MARC (BE)
FEATURES
    source
        1..5082
            Location/Qualifiers
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN

```

```

Query Match 100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 60
Db 2669 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 2728

Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 101
Db 2729 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 2769

RESULT 4
BD085110
LOCUS BD085110 5082 bp DNA linear PAT 27-AUG-2002
DEFINITION Vertebrate homologues of UNC-53 protein of C elegans.
ACCESSION BD085110
VERSION BD085110.1 GI:22630720
KEYWORDS JP 2001522222-A/8.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS Platteeuw,C.J., Arjol,C.M.B., Deraeymaeker,M., Verhasselt,P., Pujol,N.J.R., Luc, Maertens,J.S., Luyten,W., Geerts,H., Vandekerckhove,J.S., Geysen,J. and Bogaert,T.A.O.E.
TITLE Vertebrate homologues of UNC-53 protein of C elegans
JOURNAL Patent: JP 2001522222-A 8 13-NOV-2001; JANSSEN PHARMACEUTICA NV
COMMENT OS Unidentified
PN JP 2001522222-A/8
PD 13-NOV-2001
PF 03-DEC-1997 JP 1998525231
PR 04-DEC-1996 GB 9625283.8
PI CHRIST JULES PLATTEEUW,CARLOS MANUEL BUESA ARJOL,MARC PI DERAEMYAEKER,
PI PETER VERHASSELT,NATHALIE JEANNE RAYMONDE PUJOL,LUC PI JACQUES SIMON MAERTENS,
PI WALTER LUYTEN,HUGO GEERTS,JOEL STEFAAN VANDEKERCKHOVE,JOHAN PI GEYSEN,
PI THIERRY ANDRE OLIVIER EDDY BOGAERT
PC C12N15/12,C12N5/10,C12N15/85,C07K14/435,C07K16/18,A61K38/17, PC A61K49/00,
PC C12Q1/02,G01N33/53
CC Strandedness: Double;
CC Topology: Circular;
CC Vertebrate homologues of UNC-53 protein of C elegans
FT source
FT 1..5082
    Location/Qualifiers
        /organism='Unidentified'.
FEATURES
    source
        1..5082
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 60
Db 2669 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 2728

Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 101
Db 2729 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 2769

RESULT 5

```

BD234590  
LOCUS BD234590 5432 bp DNA linear PAT 17-JUL-2003  
DEFINITION Screening assay of Abeta-peptide.  
ACCESSION BD234590  
VERSION BD234590.1 GI:33044360  
KEYWORDS JP 2002531141-A/2.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 5432)  
AUTHORS Peraus,G.  
TITLE Screening assay of Abeta-peptide  
JOURNAL Patent: JP 2002531141-A 2 24-SEP-2002;  
COMMENT AVENTIS PHARMA DEUTSCHLAND GMBH  
OS Artificial Sequence  
PN JP 2002531141-A/2  
PD 24-SEP-2002  
PF 07-NOV-1999 JP 2000586944  
PR 07-DEC-1998 DE 198 56 261.6  
PT GISELA PERAUS  
PC C12N15/09,A01K67/033,A61K45/00,A61P25/28,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12Q1/37,C12Q1/68,C12N15/00,C12N5/00 CC Description  
of Artificial Sequence: Mutagen  
FH Key Location/Qualifiers  
FT source 1..5432  
FT source 1..5432 /organism='Artificial Sequence'.  
FEATURES  
source  
1..5432  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5432;  
Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 60  
Db 985 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 1044  
Qy 61 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 101  
Db 1045 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 1085  
RESULT 6  
AX026821  
LOCUS AX026821 5432 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 9 from Patent DE19856261.  
ACCESSION AX026821  
VERSION AX026821.1 GI:10187947  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Peraus,G.  
JOURNAL Patent: DE 19856261-C 9 30-MAR-2000;  
HOBCHST MARION ROUSSEL DE GMBH (DE)  
FEATURES  
source  
1..5432  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutagen"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5432;  
Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 60  
Db 985 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 1044  
Qy 61 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 101  
Db 1045 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 1085  
RESULT 7  
AX226281  
LOCUS AX226281 5650 bp DNA linear PAT 10-SEP-2001  
DEFINITION Sequence 2 from Patent WO0161024.  
ACCESSION AX226281  
VERSION AX226281.1 GI:15555545  
KEYWORDS Porcine circovirus  
SOURCE Porcine circovirus  
ORGANISM Porcine circovirus  
REFERENCE 1  
AUTHORS Palmer,K.E. and Pogue,G.P.  
TITLE Rolling circle replicon expression vectors  
JOURNAL Patent: WO 0161024-A 2 23-AUG-2001;  
Large Scale Biology Corporation (US)  
FEATURES  
source  
1..5650  
/organism="Porcine circovirus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:46221"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5650;  
Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 60  
Db 1956 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 2015  
Qy 61 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 101  
Db 2016 CAGGCATCTGTGTTGCTCCCTCCCGCTGCTTCCCTTGAC 2056  
RESULT 8  
AX202478  
LOCUS AX202478 5731 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 66 from Patent WO0152620.  
ACCESSION AX202478  
VERSION AX202478.1 GI:15392206  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.  
TITLE Methods and compositions to modulate expression in plants  
JOURNAL Patent: WO 0152620-A 66 26-JUL-2001;  
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL  
DISCOVERY, INC. (CA)  
FEATURES  
source  
1..5731  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="2C7-SID"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5731;  
Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTCTAGTTGC 60

```
Db 1701 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 1760
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 101
|||||
Db 1761 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 1801
|||||

RESULT 9
AF060226 5771 bp DNA circular SYN 14-AUG-2000
LOCUS Eukaryotic expression vector pCR3.1mbCL-XL
DEFINITION Eukaryotic expression vector pCR3.1mbCL-XL, complete sequence.
ACCESSION AF060226
VERSION AF060226.1 GI:3108232
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 5771)
Eukaryotic expression vector pCR3.1mbCL-XL
other sequences; artificial sequences; vectors.
AUTHORS Pirtskhalaishvili, G., Shurin, G.V., Gambotto, A., Esche, C., Wahl, M.,
Yurkovetsky, Z.R., Robbins, P.D. and Shurin, M.R.
TITLE Transduction of dendritic cells with Bcl-xL increases their
resistance to prostate cancer-induced apoptosis and antitumor
effect in mice
J. Immunol. 165 (4), 1956-1964 (2000)
JOURNAL
MEDLINE 20384788
PUBMED 10925278
REFERENCE
2 (bases 1 to 5771)
Gambotto, A., Pagliano, O., Shurin, M. and Robbins, P.D.
AUTHORS Direct Submision
TITLE Submitted (17-APR-1998) Vector Core Facility, University of
Pittsburgh, 300 Technology Drive, Pittsburgh, PA 15219, USA
JOURNAL
FEATURES
source
1. .5771
/organism="Eukaryotic expression vector pCR3.1mbCL-XL"
/mol_type="genomic DNA"
/db_xref="taxon:75965"
1. .596
/note="CMV"
638. .657
/note="T7; priming site also"
747. 1448
/note="BAUB/c form"
/codon_start=1
/product="murine BCL-XL"
/protein_id="AAC15799.1"
/db_xref="GI:3108233"
/translation="MSSQNRRLVVDLFLSYKLSQKYSWSQPSDVEENRTEAPEETEAE
RETSAINGNPSWHLASPAVNGATGCHSSSLDAREVTPMAVKQALREAGDEFELRYR
RAFSDLTSQLHITPGTAYQSFQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQ
VLVSRIASWATYLDHLEPWIENGWDTFVDLYGNNAASERKQERFNRWFLTGM
TVAGVLLGLSLFSRK"
1524. .1541
/note="pCR3.1 reverse priming site"
1827. .2500
/note="ColE1"
/complement(3082. .3870)
/codon_start=1
/product="neomycin/kanamycin resistance protein"
/protein_id="AAC15800.1"
/db_xref="GI:3108234"
/translation="MIFQDGLHAGSPAAMVERLFGYDWAQQTIGCSDAAVFRLSAQQR
PVLVKFDLGSALNEQDEARLSWLAATGVPCAAVLDVVTYAGRWLLLLGEVPGDOL
LSSHLAPAEKVSIMADAMRHLTDPATCFPHQAKHRIERARTMEAGLVDDQDQH
QGLAPAEFLKASMPDGEDLVVTHGDACLPTNIVNGRPSFGFDICGRGLGVADRYQD
IALATRDIAELGGEWADRLVLYGIAAPDSQRIAFYRLLDDEF"
3905. .4243
/note="SV40 promoter and origin"
/complement(4322. .5182)
/codon_start=1
/product="ampicillin resistance protein"
/protein_id="AAC15801.1"
/db_xref="GI:3108235"

Db 1701 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 1760
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 101
|||||
Db 1761 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 1801
|||||

/translation="MSIQHPRVALIPFEAFLPVAFAHPETLVKVKDAEDQLGARVGY
IEDLNSKILESFRFEERPMSTMFKVLLCGAVLSRIDAGQQLGRIIRHYNQDLVE
YSPVEKHLTDGMTVRELCSAAITMSDNTAANLLLTITIGPKELTAFLHNMGDHVTSL
DRWPELNEAIPIVDKEDTTPMVAATTLKLLTGLLTSLASRQQLIDMDEADKVGSLP
LRSLAPAGWFIADKSGAGRSRGIITAAALPGDKGPKSRIVVVIYTTGSAQTMDEERNQIA
EIGASLIKH"
5313. .5769
/note="f1"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5771;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 60
|||||
Db 1487 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 1546
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 101
|||||
Db 1547 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 1587
|||||

RESULT 10
AX195205 6050 bp DNA linear PAT 28-AUG-2001
LOCUS Sequence 10 from Patent WO0151626.
DEFINITION AX195205
ACCESSION AX195205
VERSION AX195205.1 GI:15385768
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1
Lu, X., Sun, L. and Zhang, Y.
AUTHORS Novel plasmid dna vectors
TITLE Patent: WO 0151626-A 10 19-JUL-2001;
JOURNAL ELM BIOPHARMACEUTICALS, INC. (US)
FEATURES
source
1. .6050
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="commercial plasmid"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6050;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 60
|||||
Db 3837 CTCGAGTCTAGAGCGCCGTTTAAACCGCTGATCAGCCTCGATGCGCTTCTAGTTGC 3896
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 101
|||||
Db 3897 CAGCCATCTGTTGTTGCCCTCCCGCGTGCCTTCTCTTGAC 3937
|||||

RESULT 11
AR278592 6082 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 535 from patent US 6512094.
DEFINITION AR278592
ACCESSION AR278592
VERSION AR278592.1 GI:29712838
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 6082)
Xu, J., Dillon, D. C., Mitcham, J. L., Harlocker, S. L., Jiang, Y.,
Kalos, M. D., Fanger, G. R., Retter, M. W., Stolk, J. A., Day, C. H.,
Vedvick, T. S., Carter, D., Li, S. X., Wang, A., Skeiky, Y. A. W.,
```

Hepler,W.T. and Henderson,R.A.  
 Compositions and methods for the therapy and diagnosis of prostate cancer  
 JOURNAL Patent: US 6512094-A 535 28-JAN-2003;  
 FEATURES Location/Qualifiers  
 source 1..6082  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60  
 Db 5928 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 101  
 Db 5988 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 6028

## RESULT 12

AR367288 AR367288 6082 bp DNA linear PAT 12-SEP-2003  
 LOCUS Sequence 535 from patent US 6329505.  
 DEFINITION AR367288  
 ACCESSION AR367288  
 VERSION AR367288.1 GI:34600263  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

## REFERENCE

1 (bases 1 to 6082)  
 Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Yuguu,J.,  
 Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A. and  
 Day,C.H.

## TITLE

Compositions and methods for therapy and diagnosis of prostate cancer

## JOURNAL

Patent: US 6329505-A 535 11-DEC-2001;  
 Location/Qualifiers

## FEATURES

source 1..6082  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60  
 Db 5928 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 101  
 Db 5988 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 6028

## RESULT 13

AR400324 AR400324 6082 bp DNA linear PAT 18-DEC-2003  
 LOCUS Sequence 535 from patent US 6620922.  
 DEFINITION AR400324  
 ACCESSION AR400324  
 VERSION AR400324.1 GI:40143590  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

## REFERENCE

1 (bases 1 to 6082)  
 Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
 Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,  
 Vedwick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,

Hepler,W.T. and Henderson,R.A.  
 Compositions and methods for the therapy and diagnosis of prostate cancer  
 JOURNAL Patent: US 6620922-A 535 16-SEP-2003;  
 FEATURES Location/Qualifiers  
 source 1..6082  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60  
 Db 5928 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 101  
 Db 5988 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 6028

## RESULT 14

AR405591 AR405591 6082 bp DNA linear PAT 18-DEC-2003  
 LOCUS Sequence 535 from patent US 6630305.  
 DEFINITION AR405591  
 ACCESSION AR405591  
 VERSION AR405591.1 GI:40154428  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

## REFERENCE

1 (bases 1 to 6082)  
 Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
 Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,  
 Vedwick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,  
 Hepler,W.T. and Henderson,R.A.

Compositions and methods for the therapy and diagnosis of prostate cancer

## JOURNAL

Patent: US 6630305-A 535 07-OCT-2003;  
 Location/Qualifiers

## FEATURES

source 1..6082  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60  
 Db 5928 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 101  
 Db 5988 CAGCCATCTGTTGTTGCGCCCTCCCGCTGCTTCCCTTGAC 6028

## RESULT 15

AR563971 AR563971 6082 bp DNA linear PAT 08-OCT-2004  
 LOCUS Sequence 535 from patent US 6759515.  
 DEFINITION AR563971  
 ACCESSION AR563971  
 VERSION AR563971.1 GI:53979022  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

## REFERENCE

1 (bases 1 to 6082)  
 Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
 Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A.W.,  
 Hepler, W.T. and Henderson, R.A.  
 TITLE Compositions and methods for the therapy and diagnosis of prostate  
 cancer

JOURNAL Patent: US 6759515-A 535 06-JUL-2004;

FEATURES Location/Qualifiers

source 1..6082

/organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCGCTTCTAGTTGC 60  
 Db 5928 CTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCGCTTCTAGTTGC 5987  
 Qy 61 CAGCATCTGTGTTGCGCCCTCCCGCTGCGCTTCTGAC 101  
 Db 5988 CAGCATCTGTGTTGCGCCCTCCCGCTGCGCTTCTGAC 6028

Search completed: July 14, 2005, 14:03:32  
 Job time : 758.618 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_3930\_4030

Perfect score: 101

Sequence: 1 ctcgagctctagagggcccg.....tccccgcgtctctcttgac 101

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002as:\*

7: Geneseq2002bs:\*

8: Geneseq2003as:\*

9: Geneseq2003bs:\*

10: Geneseq2003cs:\*

11: Geneseq2003ds:\*

12: Geneseq2004as:\*

13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	278	12	Adm76942 Multiple
2	101	100.0	880	13	AdS74212 Interleuk
3	101	100.0	1340	8	Acc62322 Human NOV
4	101	100.0	1353	10	Adc26320 Human NOV
5	101	100.0	1353	12	Adm35637 Novel hum
6	101	100.0	1353	12	Ado42484 Human NOV
7	101	100.0	1420	8	Acc62251 Human NOV
8	101	100.0	1461	8	Acc62237 Human NOV
9	101	100.0	1733	12	Ado42324 Human NOV
10	101	100.0	1770	10	Adj94793 Novel NOV
11	101	100.0	1772	10	Adj94791 Novel NOV
12	101	100.0	1772	10	Adj94795 Novel NOV
13	101	100.0	1782	12	Adm41037 Cycomegal
14	101	100.0	1822	10	Adc19336 cDNA enco
15	101	100.0	1822	10	Adc19334 cDNA enco
16	101	100.0	2017	8	Ada05887 Human NOV
17	101	100.0	2017	12	Adn63050 Human NOV
18	101	100.0	2017	12	Ado42504 Human NOV
19	101	100.0	2017	12	Ado42500 Human NOV
20	101	100.0	2022	12	Ado42506 Human NOV

21	101	100.0	3482	2	ADH11353	Adh11353 Vertebrat
22	101	100.0	3972	10	ADB33520	Adb33520 Plasmid p
23	101	100.0	4291	3	AAA75084	Aaa75084 Complete
24	101	100.0	5015	10	ADB33528	Adb33528 Expressio
25	101	100.0	5082	2	ADH11417	Adh11417 Plasmid p
26	101	100.0	5432	3	AAZ89476	Aaz89476 Transgeni
27	101	100.0	5650	4	AAH74866	Aah74866 Nucleocid
28	101	100.0	5650	8	ABX94356	Abx94356 Rolling c
29	101	100.0	5731	4	AAH11615	AAH11615 Six finge
30	101	100.0	5821	12	ADM97787	Adm97787 Gal4-DBD
31	101	100.0	6050	5	AAH10237	Aah10237 Commercia
32	101	100.0	6082	4	AAH93828	Aah93828 Human pro
33	101	100.0	6082	4	AAH93921	Aah93921 Human pro
34	101	100.0	6082	4	AAH85142	Aah85142 Human pro
35	101	100.0	6082	5	ACA59729	AcA59729 Prostate
36	101	100.0	6082	6	ABL95292	AbL95292 Human P51
37	101	100.0	6082	8	AAH56212	Aah56212 Human AB-
38	101	100.0	6082	8	AAH56211	Aah56211 Human AB-
39	101	100.0	6082	8	AAH56210	Aah56210 Human AB-
40	101	100.0	6082	8	ACC95456	Acc95456 Prostate
41	101	100.0	6082	10	ADB13985	Adb13985 Human pro
42	101	100.0	6082	10	ADG26401	Adg26401 Human pro
43	101	100.0	6085	8	AAH56213	Aah56213 Human AB-
44	101	100.0	6094	8	AAH56215	Aah56215 Human PSM
45	101	100.0	6097	8	AAH56214	Aah56214 Human AB-

ALIGNMENTS

RESULT 1

ADMT76942  
ID ADM76942 standard; DNA; 278 BP.

XX AC ADM76942;

XX DT 03-JUN-2004 (first entry)

XX DB Multiple cloning site of pYA3650 DNA sequence.

XX KW host-vector system; microorganism; vaccine; delivery; immunisation;  
XX KW poultry; coccidiosis; antibacterial; plasmid; vector; gene; ds.

XX OS Synthetic.

XX PN WO2004020643-A2.

XX PD 11-MAR-2004.

XX PF 29-AUG-2003; 2003WO-US026883.

XX PR 01-SEP-2002; 2002US-0407522P.

XX PA (UNIV ) UNIV WASHINGTON.

XX PI Curtiss R, Kong W;

XX DR WPI; 2004-239203/22.

XX PT New host-vector system comprising a host chromosome, and a vector, useful  
XX PT as a vaccine for immunizing a poultry, preferably chicken, against  
XX PT coccidiosis.

XX PS Example 13; Fig 39; 201pp; English.

XX CC The present invention describes a host-vector system comprising a host  
XX CC chromosome, and a vector. The host-vector system comprises: (a) a host  
XX CC chromosome comprising: (i) an activatable control sequence that is  
XX CC activatable by an inducer; (ii) a sequence that encodes a repressor,  
XX CC where the sequence is operably-linked to the activatable control sequence  
XX CC ; and (iii) at least one essential gene that encodes a polypeptide that  
XX CC is necessary for synthesis of a rigid layer of a cell wall of a  
XX CC prokaryote, and where the essential gene is inactivated; and (b) a vector

comprising: (i) a eukaryotic expression cassette comprising a eukaryotic promoter sequence, a site for insertion of a gene encoding a desired gene product and a polyadenylation sequence; (ii) a prokaryotic activator-promoter sequence; (iii) at least one origin of replication (ori); (iv) a regulatable prokaryotic promoter, which is repressible by the repressor; (v) at least one essential gene that is necessary for synthesis of a rigid layer of a cell wall of a prokaryote; (vi) at least one transcription terminator sequence; and (vii) at least one cpg sequence motif that enhances immunogenicity. Also described: (1) a microorganism comprising the host-vector system; (2) a vaccine comprising the microorganism; (3) a method for delivery of a nucleic acid vector and/or a desired gene product to a eukaryotic host; and (4) a method of immunising a poultry against coccidiosis. The host-vector system has antibacterial activity. The host vector system is useful as a vaccine for immunising a poultry against coccidiosis. The present sequence represents a nucleotide sequence which is used in the exemplification of the present invention.

XX Sequence 278 BP; 60 A; 82 C; 65 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 278;  
Best Local Similarity 100.0%; Pred. No. 8.7e-24;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60  
Db 126 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 185

Oy 61 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 101

Db 186 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 226

#### RESULT 2

AD574212  
ID AD574212 standard; DNA; 880 BP.

XX AD574212;

DT 16-DEC-2004 (first entry)

DE Interleukin-2-Fc epsilon-gamma receptor transmembrane domain fusion gene.

XX Cancer; gene therapy; vaccine; human; interleukin-2; cytokine;

KW Fc epsilon-gamma receptor; receptor; IL-2tm2; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 1..880

FT /\*tag= a

FT /product= "IL2-tm2"

FT /partial

FT /transl\_except= (pos:17..19,aa:Xaa)

FT /transl\_except= (pos:23..25,aa:Xaa)

FT /transl\_except= (pos:653..655,aa:Xaa)

FT /transl\_except= (pos:662..664,aa:Xaa)

FT /transl\_except= (pos:743..745,aa:Xaa)

FT /transl\_except= (pos:764..766,aa:Xaa)

FT /transl\_except= (pos:800..802,aa:Xaa)

FT /transl\_except= (pos:845..847,aa:Xaa)

FT /transl\_except= (pos:878..880,aa:Xaa)

FT /note= "No start or stop codon: Xaa= any amino acid"

XX WO2004080404-A2.

XX 23-SEP-2004.

XX 08-MAR-2004; 2004WO-US007012.

XX 07-MAR-2003; 2003US-0452989P.

XX (UTAH ) UNIV UTAH RES FOUND.

XX

PI Samlowski W, Adams NB, McGregor J;

XX WPI; 2004-668877/65.

DR P-PSDB; ADS74209.

XX New fusion protein comprising human interleukin-2 and a transmembrane domain of a protein and enhancing the activation of cytotoxic tumor-infiltrating lymphocytes within tumor, useful in preparing a composition for treating cancer.

PS Claim 17; SEQ ID NO 15; 64pp; English.

XX The present is that of IL-2tm2, a fusion gene comprising a human interleukin-2 (IL-2) gene joined to DNA encoding the transmembrane domain of Fc epsilon-gamma. IL-2tm2 is an example of a novel fusion gene of the invention for use in cancer gene therapy that comprises a cytokine gene and a transmembrane domain gene. It is derived from IL-2tm ADS74208 by removal of a FLAG sequence and extraneous amino acids, and was expressed from vector pcDNA3.1(+/-). The fusion protein is expressed as a membrane-bound cytokine which may be displayed on the surface of mammalian tumour cells. It is believed that by inducing expression of IL-2 on the surface of tumour cells, IL-2 will activate tumour-infiltrating lymphocytes in close proximity to tumour antigens. This activation is thought to increase activation of antigen-specific T cells and hence to result in destruction of tumour cells expressing those antigens. Murine spindle cell skin cancer RD995 cells transfected with IL-2tm fusion gene or pCMV2b (empty expression vector) were implanted subcutaneously into C3H/HEB mice. Mice implanted with RD995 cells transfected with IL-2tm fusion gene showed reduced tumour growth compared with controls.

XX Sequence 880 BP; 238 A; 214 C; 198 G; 226 T; 0 U; 4 Other;

Query Match 100.0%; Score 101; DB 13; Length 880;  
Best Local Similarity 100.0%; Pred. No. 1.2e-23;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60  
Db 632 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 691

Oy 61 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 101

Db 692 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 732

#### RESULT 3

ACC62322  
ID ACC62322 standard; cDNA; 1340 BP.

XX ACC62322;

XX 23-JUN-2003 (first entry)

XX Human NOV40b encoding cDNA SEQ ID NO:173.

XX Human; NOVX; antiatherosclerotic; hypotensive; cardiant; dermatological; anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility; haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator; neuroprotective; nootropic; antiparkinsonian; metabolic; antilipaeamic; gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma; congenital heart defect; aortic stenosis; valve disease; transplantation; tuberosus sclerosis; obesity; congenital adrenal hyperplasia; diabetes; prostate cancer; metabolic disorder; neoplasia; lymphoma; uterus cancer; fertility; haemophilia; hypercoagulation; graft versus host disease; idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia; Crohn's disease; multiple sclerosis; infectious disease; cancer; cancer-associated cachexia; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; dyslipidaemia;

XX Homo sapiens.

XX





PI Catterton E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L;  
PI Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE;  
PI Padigaru M, Patturajan M, Pena CEA, Rastelli L, Shenoy S;  
PI Shimkets RA, Spaderina SK, Spytek KA, Stone DJ, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zhong M;  
XX WPI: 2003-221607/21.  
DR P-PSDB; ADM26321.  
XX  
PT New isolated NOVX polypeptide, useful for determining the presence of, or  
PT predisposition to a disease associated with altered levels of expression  
PT of the polypeptide, and for treating or preventing cancer.  
XX  
XX Claim 20; SEQ ID NO 145; 478pp; English.  
PS  
CC The invention relates to a novel isolated NOV polypeptide. The  
CC polypeptide of the invention demonstrates cytostatic activity and may be  
CC used for determining the presence of, or predisposition to a disease  
CC associated with altered levels of expression of the polypeptide,  
CC including metabolic disorders, immune disorders, neurodegenerative  
CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases  
CC and cancer. The polypeptide may also be utilised during gene therapy  
CC procedures, vaccine development and transgenic animal production. The  
CC current sequence is that of the human NOV DNA of the invention.  
XX  
SQ Sequence 1353 BP; 324 A; 354 C; 356 G; 319 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 10; Length 1353;  
Best Local Similarity 100.0%; Pred. No. 1.3e-23;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCCGCTGTATCAGCCTCGACTGTGCTTAGTTGC 60  
Db 1187 CTCGAGTCTAGAGGCCGCGTTTAAACCCGCTGTATCAGCCTCGACTGTGCTTAGTTGC 1246  
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCCTTGAC 101  
Db 1247 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCCTTGAC 1287  
RESULT 5  
ADM35637  
ID ADM35637 standard; DNA; 1353 BP.  
XX  
AC ADM35637;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Novel human NOVX gene for treating diabetes and obesity.  
XX  
KW ds; gene; antidiabetic; anorectic; screening; insulin resistance;  
KW obesity; diabetes.  
XX  
OS Homo sapiens.  
XX  
XX WO2004013347-A2.  
XX  
PD 12-FEB-2004.  
XX  
XX 06-AUG-2003; 2003WO-US024504.  
XX  
FR 06-AUG-2002; 2002US-0401315P.  
FR 06-AUG-2002; 2002US-0401316P.  
FR 06-AUG-2002; 2002US-0401627P.  
FR 07-AUG-2002; 2002US-0401788P.  
FR 15-AUG-2002; 2002US-0403620P.  
FR 20-AUG-2002; 2002US-0404649P.  
FR 20-AUG-2002; 2002US-0404674P.  
FR 20-AUG-2002; 2002US-0404676P.  
FR 22-AUG-2002; 2002US-0405121P.  
FR 22-AUG-2002; 2002US-0405232P.  
FR 23-AUG-2002; 2002US-0405400P.  
FR 23-AUG-2002; 2002US-0405684P.

PR 23-AUG-2002; 2002US-0405687P.  
PR 23-AUG-2002; 2002US-0405698P.  
PR 26-AUG-2002; 2002US-0406353P.  
PR 27-AUG-2002; 2002US-0406130P.  
PR 27-AUG-2002; 2002US-0406131P.  
PR 03-SEP-2002; 2002US-0407919P.  
PR 09-SEP-2002; 2002US-0409366P.  
PR 31-OCT-2002; 2002US-0422756P.  
PR 02-DEC-2002; 2002US-00307817.  
XX (CURA-) CURAGEN CORP.  
XX Berghs C, Ellerman K, Guo X, Li L, Ort T, Rieger DK, Vernet CAM;  
PI Zhong M;  
XX WPI: 2004-191379/18.  
DR P-PSDB; ADM35638.  
XX  
PT New NOVX nucleic acids and polypeptides, useful in identifying compounds  
PT for treating conditions such as insulin resistance, obesity and diabetes.  
XX  
PS Claim 1; SEQ ID NO 71; 325pp; English.  
XX  
CC The invention relates to novel isolated human nucleic acid molecules and  
CC their encoded proteins designated NOVX proteins. The nucleic acids and  
CC encoded polypeptides are useful in screening for compounds useful for  
CC treating conditions such as insulin resistance, obesity or diabetes. This  
CC sequence corresponds to one of the genes of the invention.  
XX  
SQ Sequence 1353 BP; 324 A; 354 C; 356 G; 319 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 12; Length 1353;  
Best Local Similarity 100.0%; Pred. No. 1.3e-23;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCCGCTGTATCAGCCTCGACTGTGCTTAGTTGC 60  
Db 1187 CTCGAGTCTAGAGGCCGCGTTTAAACCCGCTGTATCAGCCTCGACTGTGCTTAGTTGC 1246  
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCCTTGAC 101  
Db 1247 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCCTTGAC 1287  
RESULT 6  
ADO42484  
ID ADO42484 standard; cDNA; 1353 BP.  
XX  
AC ADO42484;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Human NOVX polynucleotide #167.  
XX  
KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;  
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;  
KW scleroderma; hypertension; haemophilia;  
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;  
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;  
KW cancer-associated cachexia; multiple sclerosis; fertility.  
XX  
OS Homo sapiens.  
XX  
XX US2004058338-A1.  
XX  
PD 25-MAR-2004.  
XX  
PF 02-DEC-2002; 2002US-00307817.  
XX  
PR 03-DEC-2001; 2001US-0336881P.  
PR 05-DEC-2001; 2001US-0336820P.  
PR 07-DEC-2001; 2001US-0338285P.  
PR 07-DEC-2001; 2001US-0338318P.

PR 10-DEC-2001; 2001US-0338989P.  
 PR 10-DEC-2001; 2001US-0339022P.  
 PR 11-DEC-2001; 2001US-0339314P.  
 PR 11-DEC-2001; 2001US-0339516P.  
 PR 11-DEC-2001; 2001US-0339517P.  
 PR 11-DEC-2001; 2001US-0339611P.  
 PR 12-DEC-2001; 2001US-0340981P.  
 PR 12-DEC-2001; 2001US-0341346P.  
 PR 14-DEC-2001; 2001US-0340390P.  
 PR 14-DEC-2001; 2001US-0340440P.  
 PR 14-DEC-2001; 2001US-0340565P.  
 PR 14-DEC-2001; 2001US-0340608P.  
 PR 14-DEC-2001; 2001US-0341144P.  
 PR 17-DEC-2001; 2001US-0341477P.  
 PR 17-DEC-2001; 2001US-0341540P.  
 PR 18-DEC-2001; 2001US-0341768P.  
 PR 20-DEC-2001; 2001US-0342592P.  
 PR 31-DEC-2001; 2001US-0344903P.  
 PR 01-FEB-2002; 2002US-0353286P.  
 PR 01-FEB-2002; 2002US-0353288P.  
 PR 26-FEB-2002; 2002US-0359599P.  
 PR 26-FEB-2002; 2002US-0359626P.  
 PR 26-FEB-2002; 2002US-0359671P.  
 PR 27-FEB-2002; 2002US-0359914P.  
 PR 27-FEB-2002; 2002US-0359956P.  
 PR 28-FEB-2002; 2002US-0360924P.  
 PR 28-FEB-2002; 2002US-0360964P.  
 PR 28-FEB-2002; 2002US-0361028P.  
 PR 28-FEB-2002; 2002US-0361256P.  
 PR 28-FEB-2002; 2002US-0361264P.  
 PR 05-MAR-2002; 2002US-0361770P.  
 PR 05-MAR-2002; 2002US-0362230P.  
 PR 13-MAR-2002; 2002US-0364181P.  
 PR 13-MAR-2002; 2002US-0364238P.  
 PR 15-MAR-2002; 2002US-0364978P.  
 PR 15-MAR-2002; 2002US-0365025P.  
 PR 17-APR-2002; 2002US-0373288P.  
 PR 15-MAY-2002; 2002US-0380981P.  
 PR 16-MAY-2002; 2002US-0381004P.  
 PR 17-MAY-2002; 2002US-0381495P.  
 PR 28-MAY-2002; 2002US-0383534P.  
 PR 28-MAY-2002; 2002US-0383744P.  
 PR 29-MAY-2002; 2002US-0383829P.  
 PR 29-MAY-2002; 2002US-0384024P.  
 PR 02-JUL-2002; 2002US-0393332P.  
 PR 06-AUG-2002; 2002US-0401315P.  
 PR 07-AUG-2002; 2002US-0401788P.  
 PR 20-AUG-2002; 2002US-0404676P.  
 PR 23-AUG-2002; 2002US-0405400P.  
 PR 23-AUG-2002; 2002US-0405684P.  
 PR 23-AUG-2002; 2002US-0405687P.  
 PR 23-AUG-2002; 2002US-0405698P.  
 PR 26-AUG-2002; 2002US-0406353P.  
 (AGEE//) AGEE M L.  
 (ALSO//) ALSOBROOK J P.  
 (ANDE//) ANDERSON D W.  
 (BERG//) BERGHS C.  
 (BOLD//) BOLDOG F L.  
 (BURG//) BURGESS C E.  
 (CATT//) CATTERTON E.  
 (DIPI//) DIPPO V A.  
 (EDIN//) EDINGER S R.  
 (EISE//) EISEN A.  
 (ELLE//) ELLERMAN K.  
 (GANG//) GANGOLLI E A.  
 (GERL//) GERLACH V.  
 (GORM//) GORMAN L.  
 (ROTH//) ROTHBERG B G.  
 (GUOX//) GUO X S.  
 (HERR//) HERRMANN J L.  
 (HALV//) HALVORSEN Y.  
 (JIWW//) JI W.

PA (KEKU//) KEKUDA R.  
 PA (KHRA//) KHRAMTSOV N V.  
 PA (LARO//) LAROCHELLE W J.  
 PA (LEPL//) LEPLEY D M.  
 PA (LILL//) LI L.  
 PA (MACD//) MACDOUGALL J R.  
 PA (MILL//) MILLER C E.  
 PA (ORTT//) ORT T.  
 PA (PADI//) PADIGARU M.  
 PA (PATT//) PATTURAJAN M.  
 PA (PENA//) PENNA C E A.  
 PA (PEYM//) PEYMAN J A.  
 PA (RIEG//) RIEGER D K.  
 PA (ROTH//) ROTHENBERG M E.  
 PA (SHEN//) SHENOY S G.  
 PA (SMIT//) SMITHSON G.  
 PA (SPAD//) SPADERNA S K.  
 PA (SPYT//) SPYTEK K A.  
 PA (STON//) STONE D J.  
 PA (TAUP//) TAUPIER R J.  
 PA (VERN//) VERNET C A M.  
 PA (VOSS//) VOSS E Z.  
 PA (ZHON//) ZHONG M.  
 XX  
 PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;  
 Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;  
 Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;  
 Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khramtsov NV;  
 Larochele WJ, Lepley DM, Li L, MacDougall JR, Miller CE, Ort T;  
 Padigar M, Patturajan M, Pen CEA, Peyman JA, Rieger DK;  
 Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;  
 Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;  
 WPI: 2004-268786/25.  
 P-PSDB: ADO42485.  
 XX  
 PT New human NOVX polypeptides and nucleic acid molecules, useful for  
 diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,  
 atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or  
 scleroderma.  
 XX  
 PS Claim 20; SEQ ID NO 333; 610pp; English.  
 XX  
 CC The invention relates to human NOVX polypeptides and the polynucleotides  
 encoding them. The invention also relates to antibodies specific to the  
 NOVX polypeptides. The polypeptides, polynucleotides and antibodies are  
 useful for manufacturing a medicament for treating a syndrome associated  
 with a human disease, such as a pathology associated with the NOVX  
 polypeptide. The sequences are useful for diagnosing, treating or  
 preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,  
 diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host  
 disease, scleroderma, hypertension, haemophilia, idiopathic  
 thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,  
 obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated  
 cachexia, multiple sclerosis or fertility. The nucleic acids may be used  
 as hybridisation probes, in chromosome mapping, in tissue typing, in  
 preventive medicine or in pharmacogenomics. This sequence represents a  
 human NOVX polynucleotide of the invention.  
 XX  
 SQ Sequence 1353 BP; 324 A; 355 C; 356 G; 318 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 12; Length 1353;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-23;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTCTTAGTTGC 60  
 Db 1187 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTCTTAGTTGC 1246  
 Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTCCCTTCCTTGAC 101  
 Db 1247 CAGCCATCTGTTGTTGCCCTCCCGCTCCCTTCCTTGAC 1287

RESULT 7  
ACC62251  
ID ACC62251 standard; cDNA; 1420 BP.  
XX AC  
XX ACC62251;  
XX AC  
XX  
DT 23-JUN-2003 (first entry)  
XX  
XX  
DE Human NOV3h encoding cDNA SEQ ID NO:31.  
XX  
XX Human; NOVX; antiatherosclerotic; hypotensive; cardiac; dermatological;  
KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;  
KW haemostatic; antiinflammatory; antisthmatic; anti-HIV; immunomodulator;  
KW neuroprotective; nontropic; antiparkinsonian; metabolic; antilipaeamic;  
KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;  
KW congenital heart defect; aortic stenosis; valve disease; transplantation;  
KW tuberculous sclerosis; obesity; congenital adrenal hyperplasia; diabetes;  
KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;  
KW fertility; haemophilia; hypercoagulation; graft versus host disease;  
KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;  
KW Crohn's disease; multiple sclerosis; infectious disease; cancer;  
KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic syndrome X; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W02003023001-A2.  
PN  
XX  
XX 20-MAR-2003.  
PD  
XX  
XX  
XX 09-SEP-2002; 2002W0-US028538.  
XX  
XX 07-SEP-2001; 2001US-0318120P.  
XX  
XX 07-SEP-2001; 2001US-0318184P.  
PR  
XX 10-SEP-2001; 2001US-0318430P.  
PR  
XX 17-SEP-2001; 2001US-0322636P.  
PR  
XX 17-SEP-2001; 2001US-0322781P.  
PR  
XX 17-SEP-2001; 2001US-0322816P.  
PR  
XX 19-SEP-2001; 2001US-0322817P.  
PR  
XX 20-SEP-2001; 2001US-0323519P.  
PR  
XX 20-SEP-2001; 2001US-0323631P.  
PR  
XX 20-SEP-2001; 2001US-0323636P.  
PR  
XX 25-SEP-2001; 2001US-0324969P.  
PR  
XX 26-SEP-2001; 2001US-0325091P.  
PR  
XX 26-SEP-2001; 2001US-0325091P.  
PR  
XX 14-DEC-2001; 2001US-0341144P.  
PR  
XX 05-MAR-2002; 2002US-0359599P.  
PR  
XX 03-MAY-2002; 2002US-0361663P.  
PR  
XX 17-MAY-2002; 2002US-0377908P.  
PR  
XX 29-MAY-2002; 2002US-0381483P.  
PR  
XX 02-JUL-2002; 2002US-0383863P.  
PR  
XX 17-JUL-2002; 2002US-0393332P.  
PR  
XX 13-AUG-2002; 2002US-0403517P.  
PR  
XX 06-SEP-2002; 2002US-00236417.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Agee ML, Alsbrook JP, Anderson DW, Berghs C, Boldog FL;  
PI Burgess CE, Casman SJ, Catterton E, Chant JS, Chaudhuri A;  
PI Crabtree J, Dipippo VA, Edinger SR, Eisen AJ, Ellerman K;  
PI Gangoli EA, Gerlach VL, Giot L, Gorman L, Guo X, Gusev VV, Ji W;  
PI Kikuda R, Khramsov NV, Leach MD, Lepley DM, Li L, Liu X;  
PI Maliyankar UM, Miller CE, Ooi CE, Ort T, Padigaru M, Patturajan M;  
PI Pena CEA, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Voss EZ;  
PI Spaderna SK, Spytek KA, Taupier RJ, Twomlow N, Vernet CAM, Voss EZ;  
PI Zethusen BD, Zhong M;  
XX  
XX WPI; 2003-313241/30.  
DR  
XX P-PSDB; ABR54182.  
XX

PT Novel human proteins and nucleic acid encoding the proteins, useful for  
PT diagnosis, treatment and prevention of disorders involving the human  
XX protein or nucleic acid e.g. cardiac and neurological disorders.  
XX  
XX Claim 20; Page 111-112; 460pp; English.  
XX  
XX The present invention describes isolated human NOVX proteins, where X is  
CC 1 to 42. ACC62236 to ACC62345 encode the human NOVX proteins given in  
CC ABR54167 to ABR54276. NOVX sequences have antiatherosclerotic, cardiac,  
CC hypotensive, dermatological, anorectic, immunosuppressive, cytostatic,  
CC antidiabetic, antiinfertility, haemostatic, antiinflammatory, anti-HIV,  
CC antisthmatic, metabolic, immunomodulator, neuroprotective, nontropic,  
CC antiparkinsonian and antilipaeamic activities, and can be used in gene  
CC therapy. NOVX proteins are useful for treating or preventing a pathology  
CC associated with a NOVX protein in humans and for treating a syndrome  
CC associated with the human disease. NOVX nucleic acids, proteins and  
CC antibodies can be used in the treatment and diagnosis of cardiomyopathy,  
CC atherosclerosis, hypertension, congenital heart defects, aortic stenosis,  
CC valve disease, tuberculous sclerosis, scleroderma, obesity, transplantation,  
CC congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic  
CC disorders, neoplasm, lymphoma, uterus cancer, fertility, haemophilia,  
CC hypercoagulation, idiopathic thrombocytopenic purpura, graft versus host  
CC disease, AIDS, bronchial asthma, Crohn's disease, multiple sclerosis,  
CC infectious disease, anorexia, cancer-associated cachexia, cancer,  
CC Alzheimer's disease, Parkinson's disease, immune disorders,  
CC haematopoietic disorders, dyslipidaemias, and metabolic syndrome X.  
CC ACC62346 to ACC62465 represent PCR primers and probes for human NOVX  
CC sequences, which are used in examples from the present invention.  
CC ABR54277 represents a human trypsinogen protein given in comparison with  
CC the human NOV35b protein in the exemplification of the present invention  
XX  
XX Sequence 1420 BP; 326 A; 392 C; 338 G; 364 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 101; DB 8; Length 1420;  
Best Local Similarity 100.0%; Pred. No. 1.3e-23;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTCAGTCTAGAGGCGCGTTTAAACCCGCTGATCAGCCCTCGACTGCTCTAGTTGC 60  
Db 1252 CTCAGTCTAGAGGCGCGCGTTTAAACCCGCTGATCAGCCCTCGACTGCTCTAGTTGC 1311  
Qy 61 CAGCCATCTGTTGTCCTCCCTCCCGCTCCCTTCCTTGAC 101  
Db 1312 CAGCCATCTGTTGTCCTCCCTCCCGCTCCCTTCCTTGAC 1352  
RESULT 8  
ACC62237  
ID ACC62237 standard; cDNA; 1461 BP.  
XX AC  
XX ACC62237;  
XX  
DT 23-JUN-2003 (first entry)  
XX  
XX Human NOV1b encoding cDNA SEQ ID NO:3.  
XX  
XX Human; NOVX; antiatherosclerotic; hypotensive; cardiac; dermatological;  
KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;  
KW haemostatic; antiinflammatory; antisthmatic; anti-HIV; immunomodulator;  
KW neuroprotective; nontropic; antiparkinsonian; metabolic; antilipaeamic;  
KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;  
KW congenital heart defect; aortic stenosis; valve disease; transplantation;  
KW tuberculous sclerosis; obesity; congenital adrenal hyperplasia; diabetes;  
KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;  
KW fertility; haemophilia; hypercoagulation; graft versus host disease;  
KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;  
KW Crohn's disease; multiple sclerosis; infectious disease; cancer;  
KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic syndrome X; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX



28-FEB-2002; 2002US-0360964P.  
 28-FEB-2002; 2002US-0361028P.  
 28-FEB-2002; 2002US-0361256P.  
 28-FEB-2002; 2002US-0361264P.  
 05-MAR-2002; 2002US-0361770P.  
 05-MAR-2002; 2002US-0362230P.  
 13-MAR-2002; 2002US-0364181P.  
 13-MAR-2002; 2002US-0364238P.  
 15-MAR-2002; 2002US-0364978P.  
 15-MAR-2002; 2002US-0365025P.  
 17-APR-2002; 2002US-0373288P.  
 15-MAY-2002; 2002US-0380981P.  
 16-MAY-2002; 2002US-0381004P.  
 17-MAY-2002; 2002US-0381495P.  
 28-MAY-2002; 2002US-0383534P.  
 28-MAY-2002; 2002US-0383744P.  
 29-MAY-2002; 2002US-0383829P.  
 29-MAY-2002; 2002US-0384024P.  
 02-JUL-2002; 2002US-0393332P.  
 06-AUG-2002; 2002US-0401315P.  
 07-AUG-2002; 2002US-0401788P.  
 20-AUG-2002; 2002US-0404676P.  
 23-AUG-2002; 2002US-0405400P.  
 23-AUG-2002; 2002US-0405684P.  
 23-AUG-2002; 2002US-0405687P.  
 23-AUG-2002; 2002US-0405698P.  
 26-AUG-2002; 2002US-0406353P.  
 (AGEE//) AGEE M L.  
 (ALSO//) ALSOBROOK J P.  
 (ANDE//) ANDERSON D W.  
 (BERG//) BERGHS C.  
 (BOLD//) BOLDOG F L.  
 (BURG//) BURGESS C E.  
 (CATT//) CATTERTON E.  
 (DIPI//) DIPPO V A.  
 (EDIN//) EDINGER S R.  
 (EISE//) EISEN A.  
 (ELLE//) ELLERMAN K.  
 (GANG//) GANGOLLI E A.  
 (GERL//) GERLACH V.  
 (GORM//) GORMAN L.  
 (ROTH//) ROTHBERG B G.  
 (GUOX//) GUO X S.  
 (HERR//) HERRMANN J L.  
 (HALV//) HALVORSEN Y.  
 (JIWV//) JI W.  
 (KEKU//) KEKUDA R.  
 (KHRA//) KHRAMTSOV N V.  
 (LARO//) LAROCHELLE W J.  
 (LEPL//) LEPLEY D M.  
 (LILL//) LI L.  
 (MACD//) MACDOUGALL J R.  
 (MILL//) MILLER C E.  
 (ORTT//) ORT T.  
 (PADI//) PADIGARU M.  
 (PATT//) PATTURAJAN M.  
 (PENA//) PENNA C E A.  
 (PEYM//) PEYMAN J A.  
 (RIEG//) RIEGER D K.  
 (ROTH//) ROTHENBERG M E.  
 (SHEN//) SHENOY S G.  
 (SMIT//) SMITHSON G.  
 (SPAD//) SPADERNA S K.  
 (SPTV//) SPYTEK K A.  
 (STON//) STONE D J.  
 (TAUP//) TAUPIER R J.  
 (VERN//) VERNET C A M.  
 (VOSS//) VOSS E Z.  
 (ZHON//) ZHONG M.  
 Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL,  
 Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A,  
 Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS,  
 Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khramtsov NV,  
 Larochelle WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T,  
 Padigar M, Patturajan M, Pena CE, Peyman JA, Rieger DK,  
 Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA,  
 Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;  
 WPI: 2004-268786/25.  
 P-PSDB; ADO42325.  
 New human NOVX polypeptides and nucleic acid molecules, useful for  
 diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,  
 atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or  
 scleroderma.  
 Claim 20; SEQ ID NO 173; 610pp; English.  
 The invention relates to human NOVX polypeptides and the polynucleotides  
 encoding them. The invention also relates to antibodies specific to the  
 NOVX polypeptides. The polypeptides, polynucleotides and antibodies are  
 useful for manufacturing a medicament for treating a syndrome associated  
 with a human disease, such as a pathology associated with the NOVX  
 polypeptide. The sequences are useful for diagnosing, treating or  
 preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,  
 diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host  
 disease, scleroderma, hypertension, haemophilia, idiopathic  
 thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,  
 obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated  
 cachexia, multiple sclerosis or fertility. The nucleic acids may be used  
 as hybridisation probes, in chromosome mapping, in tissue typing, in  
 preventive medicine or in pharmacogenomics. This sequence represents a  
 human NOVX polynucleotide of the invention.  
 SQ Sequence 1733 BP; 423 A; 431 C; 460 G; 419 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 12; Length 1733;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-23;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60  
 Db 1580 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 1639  
 Qy 61 CAGCCATCTGTTGTTGCCCTCCCGGTCCTCTTGAC 101  
 Db 1640 CAGCCATCTGTTGTTGCCCTCCCGGTCCTCTTGAC 1680  
 RESULT 10  
 ADJ94793  
 ID ADJ94793 standard; DNA; 1770 BP.  
 XX  
 AC ADJ94793;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Novel NOVX gene sequence #11.  
 XX  
 KW ds; gene; antidiabetic; anorectic; cardiant; hypotensive;  
 KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;  
 KW protozoacide; nootropic; neuroprotective; antiparkinsonian;  
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;  
 KW dermatological; antiasthmatic; antilipemic; gene therapy;  
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
 KW cancer; cardiovascular disease; hypertension; atherosclerosis;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;  
 KW inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;  
 KW cell differentiation; cell proliferation; hematopoiesis; wound healing;  
 KW angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.  
 OS Homo sapiens.  
 XX



```
XX PF 05-NOV-2002; 2002WO-US035464.
XX PR 05-NOV-2001; 2001US-0338626P.
XX PR 06-NOV-2001; 2001US-0333072P.
XX PR 09-NOV-2001; 2001US-0348283P.
XX PR 15-NOV-2001; 2001US-0335610P.
XX PR 16-NOV-2001; 2001US-0338543P.
XX PR 20-NOV-2001; 2001US-0336303P.
XX PR 21-NOV-2001; 2001US-0331641P.
XX PR 21-NOV-2001; 2001US-0332152P.
XX PR 27-NOV-2001; 2001US-0333461P.
XX PR 28-NOV-2001; 2001US-0333912P.
XX PR 28-NOV-2001; 2001US-0334027P.
XX PR 29-NOV-2001; 2001US-0334300P.
XX PR 30-NOV-2001; 2001US-0334421P.
XX PR 30-NOV-2001; 2001US-0334526P.
XX PR 04-DEC-2001; 2001US-0336576P.
XX PR 04-DEC-2001; 2001US-0336664P.
XX PR 07-DEC-2001; 2001US-0338344P.
XX PR 07-DEC-2001; 2001US-0338390P.
XX PR 10-DEC-2001; 2001US-0339006P.
XX PR 10-DEC-2001; 2001US-0339008P.
XX PR 11-DEC-2001; 2001US-0339286P.
XX PR 01-FEB-2002; 2002US-0353280P.
XX PR 01-FEB-2002; 2002US-0353288P.
XX PR 04-FEB-2002; 2002US-0354392P.
XX PR 04-FEB-2002; 2002US-0354393P.
XX PR 04-FEB-2002; 2002US-0354409P.
XX PR 27-FEB-2002; 2002US-0359944P.
XX PR 27-FEB-2002; 2002US-0360148P.
XX PR 05-MAR-2002; 2002US-0361790P.
XX PR 05-MAR-2002; 2002US-0361833P.
XX PR 05-MAR-2002; 2002US-0361925P.
XX PR 05-MAR-2002; 2002US-0362230P.
XX PR 05-MAR-2002; 2002US-0362625P.
XX PR 13-MAR-2002; 2002US-0364000P.
XX PR 13-MAR-2002; 2002US-0364181P.
XX PR 13-MAR-2002; 2002US-0364182P.
XX PR 13-MAR-2002; 2002US-0364197P.
XX PR 13-MAR-2002; 2002US-0364227P.
XX PR 17-MAY-2002; 2002US-0381621P.
XX PR 28-MAY-2002; 2002US-0383675P.
XX PR 17-JUL-2002; 2002US-0396703P.
XX PR 06-AUG-2002; 2002US-0401552P.
XX PR 07-AUG-2002; 2002US-0401594P.
XX PR 07-AUG-2002; 2002US-0401787P.
XX PR 15-AUG-2002; 2002US-0403619P.
XX PR 20-AUG-2002; 2002US-0404821P.
XX PR 23-AUG-2002; 2002US-0405368P.
XX PR 23-AUG-2002; 2002US-0405496P.
XX PR 23-AUG-2002; 2002US-0405526P.
XX PR 23-AUG-2002; 2002US-0405631P.
XX PR 26-AUG-2002; 2002US-0406125P.
XX PR 04-NOV-2002; 2002US-00287226.
XX PR (CURA-) CURAGEN CORP.
XX PR Agee MD, Alsbrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;
XX PR Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
XX PR Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khrantsov NV;
XX PR Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millet I;
XX PR Ooi CE, Ort T, Padigaru M, Patturajan M, Rastelli L, Rieger DK;
XX PR Rothenberg MB, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
XX PR Vernet CAM, Zerhusen BD, Zhong M;
XX PR WPI: 2003-441551/41.
XX PR P-PSDB; ADJ94792.
XX PR
XX PR New isolated NOVX polypeptides and polynucleotides, useful for
XX PR preventing, diagnosing or treating NOVX-associated disorders, e.g.
XX PR osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
XX PR asthma, or infections.
XX PF
XX PS Claim 20; SEQ ID NO 19; 800pp; English.
XX CC The invention relates to novel isolated polypeptides, mature forms of
XX CC these, or a sequence that is at least 95 % identical to, or having one or
XX CC more conservative amino acid substitutions in the polypeptides. The
XX CC polypeptides, nucleic acid molecules and antibodies are useful in the
XX CC manufacture of a medicament for treating a syndrome associated with a
XX CC human disease, preferably a NOVX-associated disorder. The nucleic acid
XX CC molecules, polypeptides and antibodies are useful for treating,
XX CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
XX CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
XX CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
XX CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
XX CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
XX CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
XX CC dyslipidemias. The nucleic acids and polypeptides may also be used as
XX CC targets for the identification of small molecules that modulate or
XX CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation, in
XX CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
XX CC generation of antibodies that bind immunospecifically to NOVX substances
XX CC for use in therapeutic or diagnostic methods. The nucleic acids are
XX CC further used as hybridization probes, in chromosome mapping, tissue
XX CC typing, preventive medicine, and pharmacogenomics. This sequence
XX CC corresponds to the gene encoding one of the NOVX polypeptides of the
XX CC invention.
XX SQ Sequence 1772 BP; 486 A; 428 C; 427 G; 431 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 1772;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCCTCGACTGTCCTCTAGTTGC 60
Db 1606 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCCTCGACTGTCCTCTAGTTGC 1665
Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGGTCCTCTCTTGAC 101
Db 1666 CAGCCATCTGTTGTTGGCCCTCCCGGTCCTCTCTTGAC 1706
RESULT 12
ADJ94795
ID ADJ94795 standard; DNA; 1772 BP.
XX AC ADJ94795;
XX XX
XX 06-MAY-2004 (first entry)
XX DT
XX DE Novel NOVX gene sequence #12.
XX DE
XX KW ds; gene; antidiabetic; anorectic; cardiant; hypotensive;
XX KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
XX KW protozoacide; nootropic; neuroprotective; antiparkinsonian;
XX KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
XX KW dermatological; antiasthmatic; antilipemic; gene therapy;
XX KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX KW cancer; cardiovascular disease; hypertension; atherosclerosis;
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX KW epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
XX KW inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
XX KW cell differentiation; cell proliferation; hematopoiesis; wound healing;
XX KW angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO2003040325-A2.
XX XX
XX PD 15-MAY-2003.
XX XX
XX PF 05-NOV-2002; 2002WO-US035464.
XX XX
```



PR 05-NOV-2001; 2001US-0338628P.  
PR 06-NOV-2001; 2001US-0333072P.  
PR 09-NOV-2001; 2001US-0348283P.  
PR 15-NOV-2001; 2001US-0335610P.  
PR 16-NOV-2001; 2001US-0338543P.  
PR 20-NOV-2001; 2001US-0331630P.  
PR 20-NOV-2001; 2001US-0331641P.  
PR 21-NOV-2001; 2001US-0332152P.  
PR 27-NOV-2001; 2001US-0333461P.  
PR 28-NOV-2001; 2001US-0333912P.  
PR 29-NOV-2001; 2001US-0334027P.  
PR 29-NOV-2001; 2001US-0334300P.  
PR 30-NOV-2001; 2001US-0334421P.  
PR 30-NOV-2001; 2001US-0334526P.  
PR 04-DEC-2001; 2001US-0336576P.  
PR 07-DEC-2001; 2001US-0336664P.  
PR 07-DEC-2001; 2001US-0338314P.  
PR 07-DEC-2001; 2001US-0338390P.  
PR 10-DEC-2001; 2001US-0339008P.  
PR 10-DEC-2001; 2001US-0339008P.  
PR 11-DEC-2001; 2001US-0339286P.  
PR 01-FEB-2002; 2002US-0353280P.  
PR 01-FEB-2002; 2002US-0353288P.  
PR 04-FEB-2002; 2002US-0354392P.  
PR 04-FEB-2002; 2002US-0354393P.  
PR 04-FEB-2002; 2002US-0354409P.  
PR 27-FEB-2002; 2002US-0359944P.  
PR 27-FEB-2002; 2002US-0360148P.  
PR 05-MAR-2002; 2002US-0361790P.  
PR 05-MAR-2002; 2002US-0361833P.  
PR 05-MAR-2002; 2002US-0361925P.  
PR 05-MAR-2002; 2002US-0362230P.  
PR 05-MAR-2002; 2002US-0362625P.  
PR 13-MAR-2002; 2002US-0364000P.  
PR 13-MAR-2002; 2002US-0364181P.  
PR 13-MAR-2002; 2002US-0364182P.  
PR 13-MAR-2002; 2002US-0364197P.  
PR 13-MAR-2002; 2002US-0364227P.  
PR 17-MAY-2002; 2002US-0381621P.  
PR 28-MAY-2002; 2002US-0383675P.  
PR 17-JUL-2002; 2002US-0396703P.  
PR 06-AUG-2002; 2002US-0401552P.  
PR 07-AUG-2002; 2002US-0401594P.  
PR 07-AUG-2002; 2002US-0401787P.  
PR 15-AUG-2002; 2002US-0403619P.  
PR 20-AUG-2002; 2002US-0404821P.  
PR 23-AUG-2002; 2002US-0405368P.  
PR 23-AUG-2002; 2002US-0405496P.  
PR 23-AUG-2002; 2002US-0405496P.  
PR 23-AUG-2002; 2002US-0405631P.  
PR 26-AUG-2002; 2002US-0406125P.  
PR 04-NOV-2002; 2002US-00287226.  
XX (CURA-) CURAGEN CORP.  
XX  
XX  
XX Agge M., Alsobrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;  
PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;  
PI Gangolli EA, Gorman L, Gerlach VL, Kekuda R, Khrantsov NV;  
PI Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millett I;  
PI Ooi CE, Ort T, Padigaru M, Patturajan M, Rastelli L, Rieger DK;  
PI Rothenberg ME, Shenoy SG, Spaderna SK, Taupier RJ;  
PI Vernet CAM, Zerhusen BD, Zhong M;  
XX  
XX WPI; 2003-441551/41.  
DR P-PSDB; ADU94796.  
XX  
XX  
XX New isolated NOVX polypeptides and polynucleotides, useful for  
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.  
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,  
PT asthma, or infections.  
XX  
XX Claim 20; SEQ ID NO 23; 800pp; English.  
XX

CC The invention relates to novel isolated polypeptides, mature forms of  
CC these, or a sequence that is at least 95 % identical to, or having one or  
CC more conservative amino acid substitutions in the polypeptides. The  
CC polypeptides, nucleic acid molecules and antibodies are useful in the  
CC manufacture of a medicament for treating a syndrome associated with a  
CC human disease, preferably a NOVX-associated disorder. The nucleic acid  
CC molecules, polypeptides and antibodies are useful for treating.  
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and  
CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,  
CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),  
CC hematopoietic disorders, inflammatory skin disorders, asthma, and various  
CC dyslipidemias. The nucleic acids and polypeptides may also be used as  
CC targets for the identification of small molecules that modulate or  
CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation, in  
CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in  
CC generation of antibodies that bind immunospecifically to NOVX substances  
CC for use in therapeutic or diagnostic methods. The nucleic acids are  
CC further used as hybridization probes, in chromosome mapping, tissue  
CC typing, preventive medicine, and pharmacogenomics. This sequence  
CC corresponds to the gene encoding one of the NOVX polypeptides of the  
CC invention.  
XX  
SQ Sequence 1772 BP; 486 A; 428 C; 427 G; 431 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 10; Length 1772;  
Best Local Similarity 100.0%; Pred. No. 1.4e-23;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCAGTCTAGAGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGCGCTTCTAGTTGC 60  
DB 1606 CTCAGTCTAGAGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGCGCTTCTAGTTGC 1665  
QY 61 CAGCCATCTGTTGTCCTCCCTCCCGCTGCTTCTCTTGAC 101  
DB 1666 CAGCCATCTGTTGTCCTCCCTCCCGCTGCTTCTCTTGAC 1706  
RESULT 13  
AD41037, standard; DNA; 1782 BP.  
XX  
XX AD41037;  
XX  
XX 17-JUN-2004 (first entry)  
XX  
XX Cytomegalovirus nucleotide sequence SEQ ID NO:5.  
XX  
XX engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
XX Cytomegalovirus.  
XX  
XX WO2004027029-A2.  
XX  
XX 01-APR-2004.  
XX  
XX 17-SEP-2003; 2003WO-US029251.  
XX  
XX 19-SEP-2002; 2002US-0411790P.  
XX  
XX (XIME-) XIMEREX INC.  
XX  
XX Beschoner WE, Sosa CE, Thompson SC;  
XX  
XX WPI; 2004-295402/27.  
XX  
XX Engrafting foreign replacement cells within a fetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a fetal non-human mammal host.  
XX

PS Disclosure; SEQ ID NO 5; 48pp; English.

XX The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises

CC selectively destroying native cells in a tissue of a foetal non-human

CC mammal host, where the number of maternal cells of the same tissue is not

CC substantially reduced, and implanting foreign replacement cells in the

CC tissue of the foetal non-human mammal host, where the foreign replacement

CC cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to

CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present

CC sequence represents a nucleotide sequence given in the Sequence Listing

CC of the present invention but not mentioned further within the

CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 1782;

Best Local Similarity 100.0%; Pred. No. 1.4e-23;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCTAGTTGC 60

Db 1486 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCTAGTTGC 1545

Qy 61 CAGCATCTGTTGTTTGCCTCCCGCTGCTTCTTCTTGC 101

Db 1546 CAGCATCTGTTGTTTGCCTCCCGCTGCTTCTTCTTGC 1586

RESULT 14

ACD19336

ID ACD19336 standard; cDNA; 1822 BP.

XX AC ACD19336;

XX AC ACD19336;

XX 25-AUG-2003 (first entry)

DE cDNA encoding novel human protein #16.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;

KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;

KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;

KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;

KW stroke; infection; gene; ss.

OS Homo sapiens.

XX WO2003023002-A2.

PN 20-MAR-2003.

PD 09-SEP-2002; 2002WO-US028539.

PF 07-SEP-2001; 2001US-0318120P.

PR 07-SEP-2001; 2001US-0318130P.

PR 10-SEP-2001; 2001US-0318430P.

PR 17-SEP-2001; 2001US-0322636P.

PR 17-SEP-2001; 2001US-0322781P.

PR 17-SEP-2001; 2001US-0322816P.

PR 17-SEP-2001; 2001US-0322817P.

PR 19-SEP-2001; 2001US-0323519P.

PR 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001; 2001US-0323636P.

PR 25-SEP-2001; 2001US-0324969P.

PR 25-SEP-2001; 2001US-0325091P.

PR 26-SEP-2001; 2001US-0324990P.

PR 17-APR-2002; 2002US-0373212P.

PR 06-SEP-2002; 2002US-00236177.

XX (CURA-) CURAGEN CORP.

PA Spytch KA, Patturajan M, Gorman L, Li L, Anderson DM, Zhong M;

XX Gerslach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

PI Legley DM, Edinger SR, Burgess CE;

XX WPI; 2003-313242/30.

DR P-PSDB; ABO14643.

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

PT and polynucleotides, useful in gene therapy, e.g. for treating or

PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

PT stroke or infections.

PS Claim 20; Page 119-120; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX

CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated

CC with NOVX polypeptide. The DNA encoding the protein is useful in gene

CC therapy for treating the disease or condition. In particular, the NOVX

CC polypeptide or polynucleotide is useful for treating endocrine/

CC metabolism-related diseases (e.g. obesity or diabetes), central nervous

CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,

CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune

CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,

CC asthma, inflammatory bowel disease, rheumatoid arthritis or

CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,

CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver

CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),

CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).

CC These are also useful in developing powerful assay system for functional

CC analysis of various human disorders, as well as in diagnostic

CC applications, and for monitoring the effects of drugs during clinical

CC trials. This sequence encodes a novel human NOV protein

XX SQ Sequence 1822 BP; 415 A; 552 C; 500 G; 355 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 10; Length 1822;

Best Local Similarity 100.0%; Pred. No. 1.4e-23;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCTAGTTGC 60

Db 1656 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCTAGTTGC 1715

Qy 61 CAGCATCTGTTGTTTGCCTCCCGCTGCTTCTTCTTGC 101

Db 1716 CAGCATCTGTTGTTTGCCTCCCGCTGCTTCTTCTTGC 1756

RESULT 15

ACD19334

ID ACD19334 standard; cDNA; 1822 BP.

XX AC ACD19334;

XX AC ACD19334;

XX 25-AUG-2003 (first entry)

DE cDNA encoding novel human protein #14.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;  
colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;  
prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;  
lung disease; emphysema; obstructive pulmonary disease; haemophilia;  
stroke; infection; gene; ss.

Homo sapiens.

W02003023002-A2.

20-MAR-2003.

09-SEP-2002; 2002WO-US028539.

07-SEP-2001; 2001US-0318120P.

07-SEP-2001; 2001US-0318130P.

10-SEP-2001; 2001US-0318430P.

17-SEP-2001; 2001US-0322636P.

17-SEP-2001; 2001US-0322781P.

17-SEP-2001; 2001US-0322816P.

19-SEP-2001; 2001US-0322817P.

20-SEP-2001; 2001US-0323151P.

20-SEP-2001; 2001US-0323631P.

25-SEP-2001; 2001US-0324969P.

25-SEP-2001; 2001US-0325091P.

26-SEP-2001; 2001US-0324990P.

17-APR-2002; 2002US-0373212P.

06-SEP-2002; 2002US-00236177.

(CURA-) CURAGEN CORP.

Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

Rieger DK, Taupier RJ, Shenoy SG, Llu X, Padigaru M, Alsobrook JP;

Lepley DM, Edinger SR, Burgess CE;

WPI; 2003-313242/30.

P-PSDB; ABO14641.

New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

and polynucleotides, useful in gene therapy, e.g. for treating or

preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

stroke or infections.

Claim 20; Page 118-119; 586pp; English.

The invention describes a new isolated polypeptide (NOVX). The NOVX

polypeptide, nucleic acid and antibody are useful as therapeutics,

particularly in the manufacture of a medicament for treating a syndrome

associated with a human disease, which includes a pathology associated

with NOVX polypeptide. The DNA encoding the protein is useful in gene

therapy for treating the disease or condition. In particular, the NOVX

polypeptide or polynucleotide is useful for treating endocrine/

metabolism-related diseases (e.g. obesity or diabetes), central nervous

system disorders (e.g. Alzheimer's disease, Parkinson's disease,

epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune

and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,

asthma, inflammatory bowel disease, rheumatoid arthritis or

osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,

prostate or brain cancers, or melanoma), liver diseases (e.g. liver

cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),

haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).

These are also useful in developing powerful assay system for functional

analysis of various human disorders, as well as in diagnostic

applications, and for monitoring the effects of drugs during clinical

trials. This sequence encodes a novel human NOV protein

Sequence 1822 BP; 415 A; 552 C; 500 G; 355 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 10; Length 1822;

Best Local Similarity 100.0%; Pred. No. 1.4e-23;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60

Db 1656 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 1715

Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 101

Db 1716 CAGCCATCTGTTGTTGGCCCTCCCGCGTGCCTTCCTTGAC 1756

Search completed: July 14, 2005, 07:01:42

Job time : 146.448 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_3930\_4030

Perfect score: 101

Sequence: 1 ctgagctctagagggcccg.....tcccocgtgcttcttcgtac 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hlc.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_gse1.\*

9: gb\_gse2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	378	7	CF315931 HD--05-A1
2	94.6	93.7	605	7	CK719567 19817 Swo
3	77.4	76.6	400	7	CK850860 10869 Sto
4	75	74.3	295	7	CN778129 pgn2c.pk0
5	75	74.3	519	5	BM888450 TM108 Hum
6	75	74.3	521	5	BM887817 TM553 Hum
7	74	73.3	233	9	CR154962 Forward s
8	73.4	72.7	534	5	BM887701 TM304 Hum
9	73.2	72.5	132	9	CR074510 Forward s
10	73.2	72.5	329	9	CG632479 OST350781
11	72.2	71.5	286	9	CR083191 Forward s
12	70.8	70.1	130	9	BM888352 Forward s
13	69.8	69.1	75	9	CR037248 Forward s
14	69.4	68.7	600	5	BM887768 TM397 Hum
15	69.2	68.5	166	9	CR092687 Forward s
16	68.4	67.7	304	9	CR0997931 Forward s
17	67	66.3	471	4	BM819796 K-EST0087
18	66.8	66.1	87	9	CR106833 Forward s
19	66.4	65.7	104	9	CR104210 Forward s
20	66.4	65.7	110	9	BM882981 Forward s
21	66.2	65.5	284	9	BM884480 Forward s
22	66	65.3	158	9	CR117924 Forward s
23	65.8	65.1	108	9	CR173214 Forward s
24	65.6	65.0	158	9	CR018574 Forward s

c 25	65.6	65.0	197	9	CR014355	CR014355 Forward s
c 26	65	64.4	323	9	CR100521	CR100521 Forward s
c 27	64	63.4	141	9	CR126132	CR126132 Forward s
c 28	64	63.4	159	9	CR133954	CR133954 Forward s
c 29	64	63.4	160	9	CR012517	CR012517 Forward s
c 30	63.8	63.2	109	9	CR108493	CR108493 Forward s
c 31	63.4	62.8	169	9	CR137375	CR137375 Forward s
c 32	63.4	62.8	234	9	CR070494	CR070494 Forward s
c 33	63.2	62.6	89	9	CR081749	CR081749 Forward s
c 34	63.2	62.6	330	9	CR006502	CR006502 Forward s
c 35	63	62.4	106	9	CR160976	CR160976 Forward s
c 36	63	62.4	605	5	BM888562	BM888562 TMM237 Hu
c 37	62.8	62.2	77	9	CR171087	CR171087 Forward s
c 38	62.8	62.2	132	9	CR081810	CR081810 Forward s
c 39	62.8	62.2	347	9	CR045655	CR045655 Forward s
c 40	62.4	61.8	141	9	CR140069	CR140069 Forward s
c 41	62.2	61.6	107	9	CR031231	CR031231 Forward s
c 42	62.2	61.6	109	9	CR065972	CR065972 Forward s
c 43	61.6	61.0	107	9	CR093214	CR093214 Forward s
c 44	61.4	60.8	113	9	CR100912	CR100912 Forward s
c 45	61.2	60.6	115	9	CR073933	CR073933 Forward s

## ALIGNMENTS

RESULT 1  
CF315931  
LOCUS  
DEFINITION HD--05-A13.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA  
Library (HD) Oryza sativa (japonica cultivar-group) cDNA clone  
HD--05-A13, mRNA sequence.  
ACCESSION CF315931  
VERSION CF315931.1 GI:33687692  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORIGIN Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 378)  
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,  
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

## FEATURES

source  
1..378  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="HD--05-A13"  
/tissue\_type="callus"  
/dev\_stage="proliferated callus on 2N6 media for 2 weeks"  
/lab\_host="E.coli DH108"  
/clone\_lib="OSHDAC1-overexpressing transgenic rice plasmid  
cDNA library (HD)"  
/note="vector: pCR4-TOPO; Site 1: EcoRI; Callus was  
treated with ABA(20um) for 1hr. Oligo-capped mRNA was  
reverse transcribed and then used for PCR. mRNA was  
derived from rice Histone Deacetylase overexpression  
line."

## ORIGIN

Query Match 100.0%; Score 101; DB 7; Length 378;  
Best Local Similarity 100.0%; Pred. No. 5e-21;

```

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||||
Db 86 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 145
    |||||||

Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 101
    |||||||
Db 146 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 186
    |||||||

RESULT 2
CK719567
LOCUS 19817 Swollen Stolon Solanum tuberosum cDNA, mRNA sequence.
DEFINITION CK719567
ACCESSION CK719567
VERSION CK719567.1 GI:42511281
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 605)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., and
DeKoeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and
Regan,S.
TITLE Generation of ESTs from swollen stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Clones can be requested from BioAtlantech via
bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        location/Qualifiers
            1..605
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Swollen Stolon"
                /note="Vector: pBluescript II SK(+); Site 1: EcORI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a screenhouse under natural
                conditions. RNA was isolated from swollen stolon tissue,
                3-10mm in length, which was cut from the tip, to the base
                of swelling."
ORIGIN
Query Match 93.7%; Score 94.6; DB 7; Length 605;
Best Local Similarity 96.0%; Pred. No. 5.5e-19;
Matches 97; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||||
Db 400 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 459
    |||||||

Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 101
    |||||||
Db 460 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 500
    |||||||

RESULT 3
CK850860
LOCUS 10869 Stolon Solanum tuberosum cDNA, mRNA sequence.
DEFINITION CK850860
ACCESSION CK850860
VERSION CK850860.1 GI:45239470

```

```

KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 400)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De
Koeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.
TITLE Generation of ESTs from stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        location/Qualifiers
            1..400
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Stolon"
                /note="Vector: pBluescript II SK(+); Site 1: EcORI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a screenhouse under natural
                conditions. RNA was isolated from stolon tissue."
ORIGIN
Query Match 76.6%; Score 77.4; DB 7; Length 400;
Best Local Similarity 92.9%; Pred. No. 1.3e-13;
Matches 92; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||||
Db 231 CTCGAGTCTAGAGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 290
    |||||||

Qy 61 CAGCCA-TCGTGTTGTTGCCCTCCCGCTGCTTCCTT 98
    |||||||
Db 291 CAGCCACTCTGTTGTTGCCCTCCCGCTGCTTCCT 329
    |||||||

RESULT 4
CK778129
LOCUS 295 bp mRNA linear EST 20-MAY-2004
DEFINITION pgn2c.pk001.h10.f Chicken Lymphoid cDNA library (pgn2c) Gallus
gallus cDNA clone pgn2c.pk001.h10.f 3'end of pat.pk0008.d12 5',
mRNA sequence.
ACCESSION CK778129
VERSION CK778129.1 GI:47548763
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 295)
AUTHORS Morgan,R.W. and Burnside,J.
TITLE Chicken ESTs from lymphoid tissue- 3' sequence
JOURNAL Unpublished (2004)
COMMENT Contact: Robin W. Morgan
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341
Fax: 302-831-2822
Email: morgan@udel.edu, www.chickest.udel.edu.
FEATURES
    source
        location/Qualifiers
            1..295
                /organism="Gallus gallus"
                /mol_type="mRNA"

```

/db\_xref="taxon:9031"  
 /clone="pgn2c.pk001.h10.f 3'end of pat.pk0008.d12"  
 /sex="Male and Female"  
 /tissue type="thymus, bursa, spleen, PBL, bone marrow"  
 /lab host="E.coli EMDH108"  
 /clone\_lib="Chicken Lymphoid cDNA library (pgn2c)"  
 /note="Vector: pcwvSPORT 6"

## ORIGIN

Query Match 74.3%; Score 75; DB 7; Length 295;  
 Best Local Similarity 94.0%; Pred. No. 6.8e-13;  
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 78  
 |||  
 Db 1 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 60

Qy 79 CCCTCCCGCTGCTTCTCTTGAC 101  
 |||  
 Db 61 CCCTCCCGCTGCTTCTCTTGAC 83

## RESULT 5

BM888450 519 bp mRNA linear EST 08-MAR-2002  
 LOCUS TW108 Human Trabecular Meshwork cDNA library Homo sapiens cDNA  
 DEFINITION clone 104447 5', mRNA sequence.

ACCESSION BM888450

VERSION BM888450.1 GI:19272194

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 519)

Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.

Expression Profile and Genome Location of cDNA Clones from an

Infant Human Trabecular Meshwork Library

Unpublished (2002)

## JOURNAL

COMMENT Contact: Wirtz MK

Glaucoma Genetics Lab

Oregon Health Sciences University

3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA

Tel: 503-494-4698

Fax: 503-494-6875

Email: wirtzm@ohsu.edu

Seq primer: T7 Reverse.

## FEATURES

source

1..519

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="104447"

/tissue type="eye"

/cell type="trabecular meshwork"

/dev stage="2 week to 2 year old infants"

/lab\_host="TPO10P"

/clone\_lib="Human Trabecular Meshwork cDNA library"

/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human

cDNA library made from mRNA isolated from trabecular

meshwork cells established from eyes from 6 individuals,

ages 2 weeks to 2 years. Cells were harvested at passages

3 through 6. Invitrogen made a unidirectional cDNA library

from the mRNA from the frozen cells using a pCDNA3 vector

and TPO10P, host cells."

## ORIGIN

Query Match 74.3%; Score 75; DB 5; Length 519;  
 Best Local Similarity 94.0%; Pred. No. 7.5e-13;  
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 78  
 |||  
 Db 1 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 60

Db 333 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 392  
 Qy 79 CCCTCCCGCTGCTTCTCTTGAC 101  
 |||  
 Db 393 CCCTCCCGCTGCTTCTCTTGAC 415

## RESULT 6

BM887817

LOCUS

DEFINITION

BM887817

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

1 (bases 1 to 521)

Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.

Expression Profile and Genome Location of cDNA Clones from an

Infant Human Trabecular Meshwork Library

Unpublished (2002)

COMMENT Contact: Wirtz MK

Glaucoma Genetics Lab

Oregon Health Sciences University

3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA

Tel: 503-494-4698

Fax: 503-494-6875

Email: wirtzm@ohsu.edu

Seq primer: T7 Reverse

High quality sequence stop: 350.

Location/Qualifiers

1..521

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="122060"

/tissue type="eye"

/cell type="trabecular meshwork"

/dev stage="2 week to 2 year old infants"

/lab\_host="TPO10P"

/clone\_lib="Human Trabecular Meshwork cDNA library"

/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human

cDNA library made from mRNA isolated from trabecular

meshwork cells established from eyes from 6 individuals,

ages 2 weeks to 2 years. Cells were harvested at passages

3 through 6. Invitrogen made a unidirectional cDNA library

from the mRNA from the frozen cells using a pCDNA3 vector

and TPO10P, host cells."

Query Match 74.3%; Score 75; DB 5; Length 521;  
 Best Local Similarity 94.0%; Pred. No. 7.5e-13;  
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 78  
 |||  
 Db 348 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTTGC 407

Qy 79 CCCTCCCGCTGCTTCTCTTGAC 101  
 |||  
 Db 408 CCCTCCCGCTGCTTCTCTTGAC 430

## RESULT 7

CR154962/c

LOCUS

DEFINITION

CR154962

ACCESSION

VERSION

233 bp DNA linear GSS 06-JUL-2004

Forward strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP182j09, genomic survey sequence.

CR154962

CR154962.1 GI:49933807

```

KEYWORDS  GSS: genome survey sequence; MICER.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 233)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
           Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
           Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
            source
              1..233
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone="MHPP182j09"
                /clone_lib="MHPP"

ORIGIN
Query Match      73.3%; Score 74; DB 9; Length 233;
Best Local Similarity 88.9%; Pred. No. 1.3e-12;
Matches 80; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 12 AGGCGCCGTTAAACCCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGT 71
Db 223 ACGACCCCATGCATCGCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGT 164

Qy 72 TGTGTGCCCCCTCCCGCTGCTTCTTGGAC 101
Db 163 TGTGTGCCCCCTCCCGCTGCTTCTTGGAC 134

RESULT 8
BM887701
LOCUS      BM887701
DEFINITION 534 bp mRNA linear EST 08-MAR-2002
           TM304 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
           clone 107917 5', mRNA sequence.
ACCESSION  BM887701.1 GI:19271430
VERSION    BM887701
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 534)
AUTHORS    Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
TITLE      Expression Profile and Genome Location of cDNA Clones from an
           Infant Human Trabecular Meshwork Library
JOURNAL    Unpublished (2002)
COMMENT    Contact: Wirtz MK
           Glaucoma Genetics Lab
           Oregon Health Sciences University
           3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
           Tel: 503-494-4698
           Fax: 503-494-6875
           Email: wirtzm@ohsu.edu
           Seq primer: T7 Reverse.
           Location/Qualifiers
             source
               1..534
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="107917"
                 /tissue_type="eye"
                 /cell_type="trabecular meshwork"
                 /dev_stage="2 week to 2 year old infants"
                 /lab_host="T0P10F"
                 /clone_lib="Human Trabecular Meshwork cDNA library"
                 /note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
                 cDNA library made from mRNA isolated from trabecular
                 meshwork cells established from eyes from 6 individuals,

GSS: genome survey sequence; MICER.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 233)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
           Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
           Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
            source
              1..233
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone="MHPP182j09"
                /clone_lib="MHPP"

ORIGIN
Query Match      72.7%; Score 73.4; DB 5; Length 534;
Best Local Similarity 92.8%; Pred. No. 2.4e-12;
Matches 77; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGTGTTGC 78
Db 445 GCTAGAGTCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGTGTTGC 504

Qy 79 CCTTCCCCCGTGCCTTCTTGGAC 101
Db 505 CCCTCCCCCGTGCCTTCTTGGAC 527

RESULT 9
CR074510/c
LOCUS      CR074510
DEFINITION 132 bp DNA linear GSS 05-JUL-2004
           Forward strand read from insert in 3'HPRT insertion targeting and
           chromosome engineering clone MHPP255d22, genomic survey sequence.
ACCESSION  CR074510.1 GI:49808100
VERSION    CR074510
KEYWORDS   GSS; genome survey sequence; MICER.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 132)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
           Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
           Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
            source
              1..132
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone="MHPP255d22"
                /clone_lib="MHPP"

ORIGIN
Query Match      72.5%; Score 73.2; DB 9; Length 132;
Best Local Similarity 90.7%; Pred. No. 2.2e-12;
Matches 78; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 12 AGGCGCCGTTAAACCCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGT 71
Db 93 ACGACCCCATGCATCGCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCCAGCCATCTGT 34

Qy 72 TGTGTGCCCCCTCCCGCTGCTTCTCT 97
Db 33 TGTGTGCCCCCTCCCGCTGCTTCTCT 8

RESULT 10
CG632479
LOCUS      CG632479
DEFINITION 329 bp mRNA linear GSS 02-OCT-2003
           OST350781 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST350781,
           mRNA sequence.
ACCESSION  CG632479
VERSION    CG632479.1 GI:37456328
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 329)

```



## AUTHORS

Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggott, J., Beltrandelio, H., Buxton, E.C., Edwards, J., Finch, R.A., Friddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Slightenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.  
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention  
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
 Contact: Zambrowicz BP  
 OmniBank

## TITLE

Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

## COMMENT

Contact: Zambrowicz BP

Lexicon Genetics Incorporated

4000 Research Forest Drive, The Woodlands, TX 77381, USA

Email: material@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as

described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)

Class: Gene Trap.

## FEATURES

source

Location/Qualifiers

1..329

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="129Sv/Ev"

/db\_xref="taxon:10090"

/clone="OST350781"

/cell\_type="embryonic stem cell"

/clone\_lib="Mus musculus 129Sv/Ev"

## ORIGIN

Query Match 72.5%; Score 73.2; DB 9; Length 329;  
 Best Local Similarity 96.2%; Pred. No. 2.5e-12;  
 Matches 75; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

24 AACCGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGTTTGCCTC 83

185 AGCTCGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGTTTGCCTC 244

84 CCCGCTGCTCTCTTGAC 101

245 CCCGCTACCTCTCTTGAC 262

## RESULT 11

CR083191/c

LOCUS

CR083191 286 bp DNA linear GSS 05-JUL-2004

Forward strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP263n05, genomic survey sequence.

ACCESSION

CR083191.1

GI:49816780

GSS; genome survey sequence; MICER.

KEYWORDS

source

ORGANISM

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 286)

ADAMS, D.J., BIGGS, P.J., COX, A.V., DAVIES, R.M., VAN DER WEYDEN, L.,

JONKERS, J., SMITH, J., PLUMB, R.W., TAYLOR, R.G., NISHIJIMA, I., YU, Y.,

ROGERS, J. and BRADLEY, A.

Direct Submission

Submitted (20-FEB-2004)

CB10 ISA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers

1..286

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="MHP263n05"

/clone\_lib="MHP"

ORIGIN

Query Match 71.5%; Score 72.2; DB 9; Length 286;  
 Best Local Similarity 90.6%; Pred. No. 5.1e-12;  
 Matches 77; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 12 AGGCCCCGTTTAAACCCGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGT 71  
 |||||  
 Db 88 ACAGACCCCATCATCGCTGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGT 29  
 |||||  
 Qy 72 TGTTCGCCCTCCCGCTGCTTCC 96  
 |||||  
 Db 28 TGTTCGCCCTCCCGCTGCTTCC 4  
 |||||

## RESULT 12

CR083748/c

LOCUS

CR083748 130 bp DNA linear GSS 05-JUL-2004

Forward strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP2120m07, genomic survey sequence.

ACCESSION

CR083748

GI:49719810

GSS; genome survey sequence; MICER.

KEYWORDS

source

ORGANISM

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 130)

ADAMS, D.J., BIGGS, P.J., COX, A.V., DAVIES, R.M., VAN DER WEYDEN, L.,

JONKERS, J., SMITH, J., PLUMB, R.W., TAYLOR, R.G., NISHIJIMA, I., YU, Y.,

ROGERS, J. and BRADLEY, A.

Direct Submission

Submitted (20-FEB-2004)

CB10 ISA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers

1..130

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="MHP2120m07"

/clone\_lib="MHP"

ORIGIN

Query Match 70.1%; Score 70.8; DB 9; Length 130;  
 Best Local Similarity 86.7%; Pred. No. 1.2e-11;  
 Matches 78; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

12 AGGCCCCGTTTAAACCCGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGT 71

|||||

120 ACAGACCCCATCATCGCTGCTGATCAGCTCGCTGCTCTAGTTGCCAGCATCTGT 61

|||||

72 TGTTCGCCCTCCCGCTGCTTCC 101

|||||

60 TGTTCGCCCTCCCGCTGCTTCC 31

|||||

RESULT 13

CR083748/c

LOCUS

CR083748 75 bp DNA linear GSS 05-JUL-2004

Forward strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP21109, genomic survey sequence.

ACCESSION

CR083748

GI:49770303

GSS; genome survey sequence; MICER.

KEYWORDS

source

ORGANISM

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 75)

ADAMS, D.J., BIGGS, P.J., COX, A.V., DAVIES, R.M., VAN DER WEYDEN, L.,

JONKERS, J., SMITH, J., PLUMB, R.W., TAYLOR, R.G., NISHIJIMA, I., YU, Y.,

ROGERS, J. and BRADLEY, A.

Direct Submission

Submitted (20-FEB-2004)

CB10 ISA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers

1..75

/organism="Mus musculus"

/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPP31109"  
/clone\_lib="MHPP"

ORIGIN

Query Match 69.1%; Score 69.8; DB 9; Length 75;  
Best Local Similarity 97.3%; Pred. No. 2.3e-11;  
Matches 71; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 28 CGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 87  
Db 74 CGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 15

Qy 88 GTGCTTCCTTGA 100  
Db 14 GTGCTTCCTTGA 2

RESULT 14  
EN887768  
LOCUS  
DEFINITION TM397 Human Trabecular Meshwork cDNA library Homo sapiens cDNA  
ACCESSION BM887768  
VERSION BM887768  
KEYWORDS EST.  
SOURCE BM887768.1 GI:19271512  
ORGANISM Homo sapiens (human)  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 600)  
AUTHORS Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.  
TITLE Expression Profile and Genome Location of cDNA Clones from an  
JOURNAL Infant Human Trabecular Meshwork Library  
COMMENT Unpublished (2002)  
Contact: Wirtz MK  
Glaucoma Genetics Lab  
Oregon Health Sciences University  
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA  
Tel: 503-494-4698  
Fax: 503-494-6875  
Email: wirtzm@ohsu.edu  
Seq primer: 17 Reverse  
High quality sequence stop: 400.  
Location/Qualifiers  
1..600  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="119752"  
/tissue type="eye"  
/cell type="trabecular meshwork"  
/dev stage="2 week to 2 year old infants"  
/lab\_host="TPO10P."  
/clone\_lib="Human Trabecular Meshwork cDNA library"  
/note="Vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human  
cDNA library made from mRNA isolated from trabecular  
meshwork cells established from eyes from 6 individuals,  
ages 2 weeks to 2 years. Cells were harvested at passages  
3 through 6. Invitrogen made a unidirectional cDNA library  
from the mRNA from the frozen cells using a pcDNA3 vector  
and TPO10P," host cells."

ORIGIN

Query Match 68.7%; Score 69.4; DB 5; Length 600;  
Best Local Similarity 90.1%; Pred. No. 4.4e-11;  
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCGACTCGCTTCTAGTTGCCAGCCATCTGTTGTTGC 78  
Db 516 GCTAGAGCTCGCTGATCGCTCGACTGTGCTTNTAGTTGCCAGCCATCTGTTGTTGC 575

Qy 79 CCCTCCCCCGTGCCTTCCTTG 99  
Db 576 CCCTCCCCCGTGCCTTCCTTG 596

RESULT 15  
CR092687/c  
LOCUS  
DEFINITION CR092687 166 bp DNA linear GSS 05-JUL-2004  
Forward strand read from insert in 3'HPRT insertion targeting and  
chromosome engineering clone MHPP224h16, genomic survey sequence.  
ACCESSION CR092687  
VERSION CR092687.1 GI:49826516  
KEYWORDS GSS; genome survey sequence; MICER.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 166)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
FEATURES  
source  
1..166  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPP224h16"  
/clone\_lib="MHPP"

ORIGIN

Query Match 68.5%; Score 69.2; DB 9; Length 166;  
Best Local Similarity 85.6%; Pred. No. 4.1e-11;  
Matches 77; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 12 AGGCGCCGTTTAAACCGCTGATCGACTGCTGCTTCTAGTTGCCAGCCATCTGT 71  
Db 156 ACACCCCATGCTGCGCTGATCGAGTGTCTTTCTAGTTGCCAGCCATCTGT 97

Qy 72 TGTTCGCCCTCCCGCTGCTTCCTTGAC 101  
Db 96 TGTTCGCCCTCCCGCTGCTTCCTTGAC 67

Search completed: July 14, 2005, 23:23:04  
Job time : 967.667 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	101	100.0	142	6	AR356490	Sequence
C 2	101	100.0	142	6	AR538046	Sequence
C 3	101	100.0	228	6	E00019	DNA coding
C 4	101	100.0	240	1	PMOENDO	
C 5	101	100.0	251	6	E00018	DNA coding
C 6	101	100.0	251	6	I01644	Sequence 1
C 7	101	100.0	344	11	HUMUT5345	Human chr1
C 8	101	100.0	400	6	BD195256	Nucleotide
C 9	101	100.0	456	6	E00892	Synthetic D
C 10	101	100.0	456	6	E01156	DNA fragment
C 11	101	100.0	456	6	E01274	DNA encoding
C 12	101	100.0	456	6	E01302	DNA encoding
C 13	101	100.0	466	6	AX260098	Sequence
C 14	101	100.0	573	6	AX260150	Sequence
C 15	101	100.0	693	6	A43586	Sequence 11
C 16	101	100.0	693	1	AR116755	Sequence
C 17	101	100.0	998	6	AY559171	Pseudomon
C 18	101	100.0	1011	1	SMTEA0GE	X97254 S.marcesc
C 19	101	100.0	1012	2	CEC11F02	Z92776 Caenorhabdi

```
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 142)
        Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
        Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
         source
           1..142
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
DB 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 48
QY 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
DB 47 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 228)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
source
  1..228
    Location/Qualifiers
      /organism="Escherichia coli"
      /mol_type="genomic DNA"
      /db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 116
QY 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
```

```
Db 115 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 75
RESULT 4
PMOENDO/c
LOCUS PMOENDO 240 bp DNA linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform
structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J.,
Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed
in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES Location/Qualifiers
         source
           1..240
             /organism="Plasmid pMM110"
             /mol_type="genomic DNA"
             /db_xref="taxon:2599"
             /plasmid="Plasmid pMM110"
ORIGIN Unreported.
Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
DB 151 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 92
QY 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
DB 91 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 251)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;
```

```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES             Location/Qualifiers
     source          1..251
                     /organism='Escherichia coli'
                     /mol_type='genomic DNA'
                     /db_xref='taxon:562'
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 60
   |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 116
   |||||||
QY 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101
   |||||||
Db 115 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 75

RESULT 6
I01644/c
LOCUS      251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 251)
AUTHORS Gilbert,W. and Talmadge,K.
TITLE Mature protein synthesis
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
President and Fellows of Harvard College; Cambridge, MA
FEATURES             Location/Qualifiers
     source          1..251
                     /organism='unknown'
                     /mol_type='unassigned DNA'
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 60
   |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 116
   |||||||
QY 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101
   |||||||
Db 115 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 75

RESULT 7
HUMUT5345
LOCUS      344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;
microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 344)
AUTHORS Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

```

```

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAACAGAGGCGCAAAATGC
Primer B: TTCGGGAAATGTCCCGCAACC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2.
FEATURES             Location/Qualifiers
     source          1..344
                     /organism='Homo sapiens'
                     /mol_type='genomic DNA'
                     /db_xref='taxon:9606'
                     /map='8'
                     /map_224
                     /standard_name='STS UT5345'
                     /complement(202..224)
     primer_bind    36..224
     primer_bind    36..60
ORIGIN
Query Match      100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 60
   |||||||
Db 141 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAACAATAAG 200
   |||||||
QY 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101
   |||||||
Db 201 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 241

RESULT 8
BD195256/c
LOCUS      400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 400)
AUTHORS Dillon,P.J., Choi,G.H. and Welch,R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC./WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
   /organism='Unidentified'.
FEATURES
   source
     1..400
     /organism='unidentified'
     /mol_type='genomic DNA'
     /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 165 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 106
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 105 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clone=pvG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
   1..456
   /organism='synthetic construct'
   /mol_type='genomic DNA'
FEATURES
   source
     1..456
     /organism='synthetic construct'
     /mol_type='genomic DNA'

```

```

ORIGIN
/db_xref='taxon:32630'
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pvG201;
FH Key Location/Qualifiers
FH promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FT Location/Qualifiers
   1..456
   /organism='synthetic construct'
   /mol_type='genomic DNA'
   /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATGTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101

```

```
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaihara,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
PI MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
PI MATSUHARA AIZO, YANAIHARA NOBORU
PC C12P21/00, C12N15/00, (C12P21/00, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Location/Qualifiers
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta-urogastron'
FT CDS 209..439
FT /product='beta-urogastron'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
```

ORIGIN		/db_xref="taxon:7227"	
Query Match		100.0%; Score 101; DB 6; Length 466;	
Best Local Similarity		100.0%; Pred. No. 8.7e-20;	
Matches 101; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy		1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 60	
Db		280 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 221	
Qy		61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101	
Db		220 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 180	
RESULT 14			
AX260150/c		AX260150 573 bp DNA linear PAT 26-OCT-2001	
LOCUS		Sequence 112 from Patent WO0172774.	
DEFINITION		AX260150	
ACCESSION		AX260150.1 GI:16509172	
VERSION			
KEYWORDS		Drosophila melanogaster (fruit fly)	
SOURCE		Drosophila melanogaster	
ORGANISM		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE		1	
AUTHORS		Deak, P., Glover, D.M. and Midgley, C.	
TITLE		Cell cycle progression proteins	
JOURNAL		Patent: WO 0172774-A 112 04-OCT-2001;	
CYCLACEI		Limited (GB)	
FEATURES		Location/Qualifiers	
source		1..573	
		/organism="Drosophila melanogaster"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:7227"	
ORIGIN			
Query Match		100.0%; Score 101; DB 6; Length 573;	
Best Local Similarity		100.0%; Pred. No. 8.8e-20;	
Matches 101; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy		1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 60	
Db		355 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 296	
Qy		61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101	
Db		295 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 255	
RESULT 15			
A43586		A43586 693 bp DNA linear PAT 06-MAR-1997	
LOCUS		Sequence 11 from Patent WO9507357.	
DEFINITION		A43586	
ACCESSION		A43586.1 GI:2298779	
VERSION			
KEYWORDS		Cuphea lanceolata	
SOURCE		Cuphea lanceolata	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Lythraceae; Cuphea.	
REFERENCE		1 (bases 1 to 693)	
AUTHORS		Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,	
		Hoerhke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,	
		Schulte, W., Voetz, M., Walek, J. and Scheil, J.	
PROMOTERS			
TITLE		Patent: WO 9507357-A 11 16-MAR-1995;	
JOURNAL		MAX PLANCK GESELLSCHAFT (DE)	
COMMENT		Other publication CA 2169093 950316	

FEATURES		Other publication AU 7615494 950327.	
source		Location/Qualifiers	
		1..693	
		/organism="Cuphea lanceolata"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3930"	
		/clone="CLKASIG8"	
		/clone_lib="Genomic Lambda Fix II"	
ORIGIN			
Query Match		100.0%; Score 101; DB 6; Length 693;	
Best Local Similarity		100.0%; Pred. No. 8.8e-20;	
Matches 101; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy		1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 60	
Db		592 AGGGTTATTGCTCATGAGCGGATACATATTGATGTTATTAGAAAATAAACAAATAG 651	
Qy		61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101	
Db		652 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 692	
Search completed:		July 14, 2005, 14:03:32	
Job time :		756.618 secs	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_7860\_7960  
Perfect score: 101  
Sequence: 1 aggttattgtctcatgacg.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues.

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	142	2	AAV76919
C 2	101	100.0	228	1	Aan10032 Sequence
C 3	101	100.0	251	1	Aan10031 Sequence
C 4	101	100.0	400	2	Aav31229 E. coli J
C 5	101	100.0	456	1	Aan60824 Plasmid p
C 6	101	100.0	456	1	Aan71080 Sequence
C 7	101	100.0	456	1	Aan70833 Beta-urog
C 8	101	100.0	456	1	Aan81765 Sequence
C 9	101	100.0	466	6	ABA90413 Drosophil
C 10	101	100.0	487	2	Aax21173 Polynucle
C 11	101	100.0	535	2	Aax21149 Polynucle
C 12	101	100.0	573	6	ABA90456 Drosophil
C 13	101	100.0	605	12	ADH58311 Electroph
C 14	101	100.0	776	4	AAS30560 DNA encod
C 15	101	100.0	776	4	AAS27819 DNA encod
C 16	101	100.0	776	4	ABK42984 Genomic s
C 17	101	100.0	776	4	Aal07344 Human rep
C 18	101	100.0	776	4	Aal03229 Human rep
C 19	101	100.0	776	4	AAL06588 Human rep
C 20	101	100.0	776	4	AAL07340 Human rep

C 21	101	100.0	776	5	ABAI4573	Abal4573 Human ner
C 22	101	100.0	776	5	AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8	ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8	ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8	ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9	ADB91869	Adb91869 Human sec
C 27	101	100.0	776	9	ADB61140	Adb61140 Connectiv
C 28	101	100.0	776	10	ADB94622	Adb94622 Novel hum
C 29	101	100.0	776	10	ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10	ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12	ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4	AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4	AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4	ABK42983	Abk42983 Genomic s
C 35	101	100.0	845	4	AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4	AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4	AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4	AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4	AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4	AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4	AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4	AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5	ABA14572	Abal4572 Human ner
C 44	101	100.0	845	5	AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9	ADB61139	Adb61139 Connectiv

ALIGNMENTS

RESULT 1  
AAV76919/c  
ID AAV76919 standard; DNA; 142 BP.  
XX  
AC AAV76919;  
XX  
DT 16-MAR-1999 (first entry)  
XX  
DE Staphylococcus aureus contig SEQ ID #2608.  
XX  
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.  
XX  
OS Staphylococcus aureus.  
XX  
FN EP786519-A2.  
XX  
PD 30-JUL-1997.  
XX  
PF 07-JAN-1997; 97EP-00100117.  
XX  
PR 05-JAN-1996; 96US-0009861P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Kunsch CA, Choi GH, Barash SC, Dillon PU, Fannon MR, Rosen CA;  
XX  
XX WPI; 1997-374922/35.  
XX  
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
PT stored on computer readable medium and used in the production of anti-  
PT S.aureus vaccines.  
XX  
PS Claim 1; Page 2287; 3271pp; English.  
XX  
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or

industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against *S. aureus* infection. The polypeptides can also be used in a kit for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium.

	Query Match	100.0%	Score 101;	DB 2;	Length 142;
	Best Local Similarity	100.0%;	Pred. No. 2.1e-21;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGGTTATTGTCATGAGCGGATACATATTTTGAATGTTATTAGAAAAATAACAATAG	60		
Db	107	AGGGTTATTGTCATGAGCGGATACATATTTTGAATGTTATTAGAAAAATAACAATAG	48		
Qy	61	GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC	101		
Db	47	GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC	7		

RESULT 2  
 AAN10032/c  
 ID AAN10032 standard; DNA; 228 BP.  
 XX  
 XX  
 AC AAN10032;  
 XX  
 XX  
 DT 13-AUG-1992 (first entry)  
 XX  
 DE Sequence of the pKT218 EcoRI-PetI penicillinase gene fragment.

Key	Location/Qualifiers
misc_feature	1. .4
	/tag= a
	/label= sticky end
misc_feature	225. .228
	/tag= b
	/label= sticky end

XX	EP38182-A.
PN	
XX	
XX	
PD	21-OCT-1981.
XX	
PF	09-APR-1981; 81EP-00301561.
XX	
PR	11-APR-1980; 80US-00139225.
XX	
PA	(HARD ) HARVARD COLLEGE.

XX  
PI  
PI  
XX  
DR  
DR

XX Synthesis of mature protein or polypeptide - by using bacterial host  
PT transformed by cloned vehicle contg. DNA fragment etc.  
XX  
PS Example: Fig 3: 34pp: English.

CC The closest identifiable promoter for the penicillinase gene in pKT241  
CC (AAN10031) is located in the region 14 to 20 nucleotides before its  
AA

translational start signal. In the examples, the 3' end of pKT241 was attached to the signal DNA sequence of the DNA fragment (19) for rat preproinsulin (see AAN10033). The closest identifiable promoter for the penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20 nucleotides before its translational start signal. In the examples, the 3' end of pKT218 was attached to the signal DNA sequence of the DNA fragment (CB6) for rat preproinsulin (see AAN10034).

	Query Match	100.0%;	Score 101;	DB 1;	Length 228;
	Best Local Similarity	100.0%;	Pred. No. 2.3e-21;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG	60		
Db	175	AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG	116		
Qy	61	GGGTTTCGGGCACATTTCCCGAAAAAGTGCACCTGACGTC	101		
Db	115	GGGTTTCGGGCACATTTCCCGAAAAAGTGCACCTGACGTC	75		

RESULT 3  
AAN10031/c  
ID AAN10031 standard; DNA; 251 BP.

AA  
AC AAN10031;

13-AUG-1992 (first entry)

DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

AA Cloning vehicle; bacterial vector; transformed host; penicillinase;  
KW insulin; ds.  
KW

XX OS Escherichia coli.

XX	key	Location/Qualifiers
FT	misc_feature	1..4
FT		/tag= a
FT		/label= sticky end
FT	misc_feature	248..251
FT		/tag= b
FT		/label= sticky end

XX	EP38182-A.	81EP-00301561.
PN		
XX		
XX	21-OCT-1981.	80US-00139225.
PD		
XX		
PF	09-APR-1981;	
XX		
PR	11-APR-1980;	

AA  
PA (HARD ) HARVARD COLLEGE.

XX PI Gilbert W. Talmadge K:

XX  
DR WPI; 1981-80125D/44.  
DR P-PSDB: AAP10038.

XX Synthesis of mature protein or polypeptide - by using bacterial host  
PT transformed by cloned vehicle contg. DNA fragment etc.  
PT

XX Example: Fig 2: 34pp: English. PS

XX CC The closest identifiable promoter for the penicillinase gene in pKT241  
CC (AAAN10031) is located in the region 14 to 20 nucleotides before its  
CC translational start signal. In the examples, the 3' end of pKT241 was  
CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
CC preproinsulin (see AAAN10033). The closest identifiable promoter for the  
CC penicillinase gene in pKT218 (AAAN10032) is located in the region 14 to 20  
CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pK218 was attached to the signal DNA sequence of the DNA  
CC fragment (CB6) for rat preproinsulin (see AAN10034)

XX Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60  
Db |||||  
175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 116  
Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101  
Db |||||  
115 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75

## RESULT 4

AAV31229/c  
ID AAV31229 standard; DNA; 400 BP.

XX AC AAV31229;

DT 01-OCT-1998 (first entry)

XX E. coli J96 pathogenicity island contig #43.

XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;  
KW PAI V; pheV; vaccine; protective immune response; ds.

XX Escherichia coli.

XX WO9822575-A2.

XX 28-MAY-1998.

XX 21-NOV-1997; 97WO-US021347.

XX 22-NOV-1996; 96US-0031626P.

PR 14-OCT-1997; 97US-0061953P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (UYWI-) UNIV WISCONSIN.

XX Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to develop  
PT products for the detection of pathogenic E. coli and to elicit an immune  
PT response to pathogenic E. coli.

PS Claim 21; Page 140-141; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing  
CC pathogenicity island (PAI) sequences, and represents a nucleic acid  
CC molecule of the invention. PAIs are large fragments of DNA which comprise  
CC pathogenicity determinants. The sequences of the invention are taken from  
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)  
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
CC approximately 94 min (at pheR) on the E. coli chromosome and is  
CC approximately 160 kb in size. Antibodies specific to the proteins encoded  
CC by the PAI open reading frames of the invention can be used in kits to  
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
CC a protective immune response in an animal to the uropathogenic E. coli  
CC strain J96

XX Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60  
Db |||||  
165 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 106

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101  
Db |||||  
105 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 65

## RESULT 5

AAN60624/c  
ID AAN60624 standard; DNA; 456 BP.

XX AC AAN60624;

XX 25-MAR-2003 (revised)

DT 29-OCT-1991 (first entry)

XX Plasmid pUG201 sequence encoding beta-urogastrone.

XX Beta-lactamase signal peptide; pGH54; pGH55; ss.

XX Synthetic.

XX Key Location/Qualifiers

FT promoter 125..170

FT /\*tag= a

FT RBS 200..203

FT /\*tag= b

FT CDS 209..439

FT /\*tag= c

FT sig\_peptide 209..277

FT /\*tag= d

FT /label= Beta-lactamase signal peptide

FT mat\_peptide 278..436

FT /\*tag= e

FT /label= Beta-urogastrone

XX WO8603779-A.

XX 03-JUL-1986.

XX 19-DEC-1985; 85WO-JP000696.

XX 21-DEC-1984; 84JP-00271206.

XX (EART ) EARTH CHEM CO LTD.

PA (OHGA/) OHGAI H.

XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

XX WPI; 1986-182911/28.

DR P-PSDB; AAP60678.

XX Recombinant vector for polypeptide secretion - contains signal peptide  
PT sequence directly bonded to peptide-coding sequence.

XX Disclosure; Table 4; 79pp; Japanese.

XX The plasmid produces secreted beta-urogastrone in a transformed  
CC expression system. Similar plasmids may be constructed where the  
CC secretion signal may be coupled with eg. somatostatin, insulin, growth  
CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,  
CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to  
CC correct PA field.)

XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60  
 |||||  
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101  
 |||||  
 Db 113 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 73  
 |||||

## RESULT 6

AAN71080/c  
 ID AAN71080 standard; DNA; 456 BP.  
 XX  
 AC AAN71080;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 13-MAY-1991 (first entry)  
 XX  
 DE Sequence encoding beta-urogastrone.  
 XX  
 KW pUGT 150s; beta-UG; ds.  
 XX  
 OS Escherichia coli.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT promoter 125..170  
 FT /\*tag= a  
 FT CDS 209..439  
 FT /\*tag= b  
 FT /transl\_except= (pos:434..436,aa:Arg)  
 FT  
 XX JPG2190083-A.  
 XX  
 XX 20-AUG-1987.  
 XX  
 PF 14-FEB-1986; 86JP-00031415.  
 XX  
 PR 14-FEB-1986; 86JP-00031415.  
 XX  
 XX (EART ) EARTH SEIYAKU KK.  
 XX  
 XX WPI; 1987-273761/39.  
 XX  
 XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.  
 XX  
 XX Disclosure; Page 553; 34pp; Japanese.  
 XX  
 CC Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing sequences comprising a tac promoter, SD site, signal peptide, and coding sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60  
 |||||  
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101  
 |||||  
 Db 113 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 73  
 |||||

## RESULT 7

AAN70833/c  
 ID AAN70833 standard; DNA; 456 BP.  
 XX  
 AC AAN70833;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 18-JAN-1991 (first entry)  
 XX  
 DE Beta-urogastrone sequence.  
 XX  
 KW Tumour; inosine; DNA probe; ds.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT promoter 125..170  
 FT /\*tag= b  
 FT RBS 200..204  
 FT /\*tag= c  
 FT CDS 209..439  
 FT /\*tag= a  
 FT sig\_peptide 209..277  
 FT /\*tag= d  
 XX  
 XX JP62244398-A.  
 XX  
 XX 24-OCT-1987.  
 XX  
 PF 16-APR-1986; 86JP-00087368.  
 XX  
 PR 16-APR-1986; 86JP-00087368.  
 XX  
 XX (SEKI ) SEKISUI CHEM IND CO LTD.  
 XX  
 XX WPI; 1987-339045/48.  
 XX  
 DR P-PSDB; AAP70505.  
 XX  
 XX Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.  
 XX  
 XX Disclosure; Page 11; 11pp; Japanese.  
 XX  
 CC An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The ssDNA and probe are hybridized and the existence of DNA in the product is detected. It can be used to detect the presence of malignant tumour.  
 CC  
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60  
 |||||  
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101  
 |||||  
 Db 113 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 73  
 |||||

## RESULT 8

AAN81765/c  
 ID AAN81765 standard; DNA; 456 BP.  
 XX  
 AC AAN81765;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 13-DEC-1990 (first entry)  
 XX

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),  
DE Arg (53).  
XX Gastric acid secretion; cell proliferation; hormone; ds.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH CDS 209..277  
FT /\*tag= a  
FT CDS 278..439  
FT /\*tag= b  
FT /product= "New beta-urogastrone deriv."  
XX JP63012298-A.  
XX 19-JAN-1988.  
XX 30-JUN-1986; 86JP-00153783.  
XX 30-JUN-1986; 86JP-00153783.  
XX (EART ) EARTH SEIYAKU KK.  
XX WPI; 1988-054638/08.  
DR P-PSDB; AAP81349.  
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and  
PT proliferation promotion activity.  
XX Disclosure; Page 685; 76pp; Japanese.  
XX The deriv. has various biological activities such as gastric acid  
CC secretion inhibiting action, or cell proliferation promoting action. The  
CC deriv. has the same biological or pharmacological activities as beta-  
CC urogastrone. It is not susceptible to denaturation by oxidn. and is  
CC chemically stable. Deriv. has resistance to proteolytic enzymes such as  
CC protease. (Updated on 25-MAR-2003 to correct PA field.)  
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 114  
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
RESULT 9  
IDA90413/c  
ID ABA90413 standard; DNA; 466 BP.  
XX AC ABA90413;  
XX 12-FEB-2002 (first entry)  
XX Drosophila cell cycle progression protein coding sequence #48.  
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;  
KW antinflammatory; antiparasitic; dermatologic; antifungal; mitosis;  
KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;  
KW cell cycle progression protein; tumour; proliferative disorder;  
KW cardiovascular; autoimmune; dermatological disorder; ds.  
XX Drosophila sp.  
XX Fraser CM;

PN WO200172774-A2.  
XX 04-OCT-2001.  
XX 23-MAR-2001; 2001WO-GB001297.  
XX 24-MAR-2000; 2000GB-00007268.  
XX (CYCL-) CYCLACEL LTD.  
XX Deak P, Glover DM, Midgley C;  
XX WPI; 2002-055132/07.  
XX Polynucleotides encoding cell cycle progression proteins, useful for  
PT treating a tumor or a proliferative disorder.  
XX Claim 1; Page 99; 213pp; English.  
XX The present invention relates to Drosophila cell cycle progression  
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-  
CC ABA90520). The coding sequences and proteins are useful for identifying a  
CC substance capable of affecting the function of the corresponding gene, a  
CC substance capable of inhibiting the cell division cycle, or capable of  
CC inhibiting mitosis and/or meiosis. They can also be used in a method for  
CC treating a tumour or proliferative disorder, cardiovascular disorders  
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as  
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders  
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic  
CC disorders (such as malaria)  
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;  
SQ Query Match 100.0%; Score 101; DB 6; Length 466;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60  
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 221  
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
Db 220 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180  
RESULT 10  
AAX21173/c  
ID AAX21173 standard; DNA; 487 BP.  
XX AC AAX21173;  
XX 05-MAY-1999 (first entry)  
XX Polynucleotide sequence from the genome of Treponema pallidum.  
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;  
KW enzyme production; ds.  
XX Treponema pallidum.  
XX WO9859034-A2.  
XX 30-DEC-1998.  
XX 23-JUN-1998; 98WO-US013041.  
XX 24-JUN-1997; 97US-0050667P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Fraser CM;





PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 08-SEP-2000; 2000US-0232082P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 03-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-476223/51.

DR Novel isolated prostate gland related polypeptide useful for diagnosis  
PT and treatment of disorders of prostate such as prostatodystonia,  
PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.  
XX Claim 1; SEQ ID NO 418; 512pp; English.  
PS The invention relates to novel isolated prostate gland related nucleic  
XX acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,  
CC prognosis, prevention, and/or treatment of diseases and/or disorders of  
CC the prostate such as acute non-bacterial prostatitis, chronic non-  
CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,  
CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic  
CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including  
CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and  
CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for  
CC diagnosing and treating reproductive system disorders (Paget's disease),  
CC autoimmune disorders (systemic lupus erythematosus, rheumatoid  
CC arthritis), blood-related disorders (sickle cell anaemia),  
CC hyperproliferative disorders, urinary system disorders  
CC (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory  
CC disorders, musculoskeletal system disorders, neural activity and  
CC neurological disorders (Alzheimer's disease and Parkinson's disease),  
CC endocrine disorders (Addison's disease), gastrointestinal disorders  
CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),  
CC pancreatic and gall bladder disorders, disorders of the large intestine,  
CC developmental and inherited disorders, diseases at the cellular level,  
CC and wound healing and epithelial cell proliferation. (I) or (II) is  
CC useful to prevent skin aging, for preventing hair loss, to maintain  
CC organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;  
Best Local Similarity 100.0%; Pred. No. 2.9e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTCGAATGTTAGAAAAATAACAATAG 60  
|||  
Db 546 AGGGTTATTGTCATGAGCGGATACATATTTCGAATGTTAGAAAAATAACAATAG 487  
|||  
Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101  
|||  
Db 486 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 446  
|||

RESULT 15  
AAS27819/c  
ID AAS27819 standard; DNA; 776 BP.





PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI  
 XX WPI; 2001-465460/50.  
 XX  
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders.  
 XX  
 XX Claim 1; SEQ ID NO 1479; 880pp; English.  
 XX  
 XX The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative disorders  
 CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
 CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
 CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
 CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
 CC respiratory disorders, dermatological disorders, in wound healing,  
 CC epithelial cell proliferation, endocrine disorders (e.g. Addison's  
 CC disease), reproductive system disorders, gastrointestinal disorder  
 CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
 CC B-cell responsiveness to pathogens, activators of T-cells, to induce  
 CC higher affinity antibodies, and as a means to induce tumour proliferation  
 CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-  
 CC AAS27850 represent novel signal transduction pathway protein coding  
 CC sequences and PCR primers of the invention  
 XX  
 XX Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;  
 SQ  
 Query Match 100.0%; Score 101; DB 4; Length 776;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTCCTCATGAGCGGATACATATTTGAATGTATTTCAGAAAAATAACAAATAG 60  
 Db 546 AGGGTTATTCCTCATGAGCGGATACATATTTGAATGTATTTCAGAAAAATAACAAATAG 487  
 QY 61 GGGTTCGCGACATATTTCCCGAAAAAGTCCACTGCAGTC 101  
 Db 486 GGGTTCGCGACATATTTCCCGAAAAAGTCCACTGCAGTC 446

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-44\_COPY\_7860\_7960  
Perfect score: 101  
Sequence: 1 agggttattgtctcatgagc.....gaaagtgcacatgcagtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 34239544 segs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	AL000426
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

#### ALIGNMENTS

RESULT 1  
BM078095/c  
LOCUS 300 bp mRNA linear EST 30-NOV-2001  
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma cylindrosporum cDNA 5', mRNA sequence.  
ACCESSION BM078095  
VERSION BM078095.1 GI:17157967  
KEYWORDS EST.  
SOURCE Hebeloma cylindrosporum  
ORGANISM Hebeloma cylindrosporum  
REFERENCE 1 (bases 1 to 300)  
AUTHORS Wipf D., Benjdia, M., Tegeder, M. and Frommer, W.B.  
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wipf D  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: pDR196 5' primer (PMA 5')  
High quality sequence stop: 300  
POLYA-No.

FEATURES  
source  
1..300  
Location/Qualifiers  
/organism="Hebeloma cylindrosporum"  
/mol\_type="mRNA"  
/strain="H1"  
/db\_xref="taxon:76867"  
/tissue\_type="Mycelia"  
/lab\_host="E. coli XLI-Blue"  
/clone\_lib="Hebeloma cylindrosporum functional cDNA library"  
/note="vector: pDR 196 (unpublished); Site\_1: EcoRI; Site\_2: XhoI"

#### ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 8.1e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGTTATTGCTCATGCGGATACATATTGTAATGCTATTAGAAAAATAACAATAG 60

```

|||||
174 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 115
|||||
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 111
|||||

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 111
|||||

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGGCGCACATTTCCCGGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139

RESULT 5
AL597149
LOCUS
DEFINITION
DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DKFZp313J1611_5', mRNA sequence.
ACCESSION
AL597149
VERSION
AL597149.1 GI:15154845
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 391)
Koehler K., Beyer A., Mewes W., Weil B. and Wiemann S.
EST (Koehler K., Beyer A., Mewes W., Weil B. and Wiemann S.)
UNPUBLISHED (1999)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139

RESULT 6
CC819240
LOCUS
DEFINITION
10005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 414)
Dunn D., Doak T., Herrick G. and Weiss R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
UNPUBLISHED (2003)
CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60
Db 414 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 355

RESULT 6
CC819240/c
LOCUS
DEFINITION
10005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 414)
Dunn D., Doak T., Herrick G. and Weiss R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
UNPUBLISHED (2003)
CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60
Db 414 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 355

RESULT 6
CC819240/c
LOCUS
DEFINITION
10005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 414)
Dunn D., Doak T., Herrick G. and Weiss R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
UNPUBLISHED (2003)
CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 60
Db 414 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAACAAATAG 355

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCST library Haplochromis chilotes cDNA clone no90c12,
            mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
        ORGANISM Haplochromis chilotes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
            Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
AUTHORS Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
TITLE Orf sequences of cichlid in Lake Victoria are essentially same
JOURNAL Unpublished (2004)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source           1..417
                     /organism="Haplochromis chilotes"
                     /mol_type="mRNA"
                     /db_xref="taxon:257977"
                     /clone="no90c12"
                     /tissue_type="jaw"
                     /dev_stage="varied"
                     /clone_lib="HCEST library"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 60
    |||||
Db 129 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 70

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 69 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 29

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10 library Sterkiella
            histriomuscorum genomic clone UUGC10006J13 R, genomic survey
            sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
        ORGANISM Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
JOURNAL macronuclear chromosomes
COMMENT Unpublished (2003)
        Contact: Robert B. Weiss
        University of Utah Genome Center
        University of Utah
        Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
        84112, USA
        Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACACGATGAC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source           1..491
                     /organism="Sterkiella histriomuscorum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:94289"
                     /clone="UUGC10006J13"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Oxytricha plasmid UUGC10 library"
                     /notes="Vector: FWD42nv; Purified macronuclear chromosomal
                     DNA from Oxytricha trifallax was blunt end-repaired with
                     T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                     oligonucleotides were ligated to the blunt ends in high
                     molar excess. Vector DNA was prepared from a derivative of
                     PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                     derivative of plasmid R1. The vector was ligated with
                     adaptors complementary to the insert adaptors and
                     purified. The sheared, adapted mouse DNA was annealed to
                     adapted vector DNA, and transformed into
                     chemically-competent E. Coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 60
    |||||
Db 412 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 353

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 352 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 312

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
            sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
        ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 495)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
            Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
        Bioinformatics and Gene Network Research Group
        Zhejiang University
        Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
        Tel: 0086-571-86892051
        Fax: 0086-571-86961525
        Email: webmaster@estarray.org, URL: http://www.estarray.org
        Seq primer: M13 forward primer.
FEATURES             Location/Qualifiers
     source           1..495
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:4530"
                     /clone="S035A01"

```

```
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
    |||||||
DB 62 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 121
    |||||||

QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 101
    |||||||
DB 122 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 162
    |||||||

RESULT 10
CC181374/c
LOCUS
DEFINITION
100004807R Oxytricha plasmid UUGC100004B07 R, genomic survey
sequence.
ACCESSION
CC181374.1 GI:32897661
VERSION
CC181374
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 495)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
FEATURES
Location/Qualifiers
1..495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004B07"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 496;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
    |||||||
DB 391 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 332
    |||||||

QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 101
    |||||||
DB 332 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 292
    |||||||

RESULT 11
CC181523/c
LOCUS
DEFINITION
100004113R Oxytricha plasmid UUGC100004L13 R, genomic survey
sequence.
ACCESSION
CC181523
VERSION
CC181523.1 GI:32897943
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 496)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
FEATURES
Location/Qualifiers
1..496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 496;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
    |||||||
DB 391 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 332
    |||||||

QY 61 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 101
    |||||||
DB 332 GGGTTCCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 292
    |||||||
```





REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Stichotrichida, Oxytrichidae; Sterkiella.  
1 (bases 1 to 518)  
Dunn,D., Doak,T., Herrick,G. and Weiss,R.  
Paired end reads from plasmid inserts of Oxytricha trifallax  
macronuclear chromosomes  
Unpublished (2003)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Plate: 0002 row: D column: 21  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 518.

FEATURES  
source  
1..518  
/organism="Sterkiella histriomuscorum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:94289"  
/clone="UUGC100002J19"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Oxytricha plasmid UUGC10 library"  
/note="Vector: PWD42nv; Purified macronuclear chromosomal  
DNA from Oxytricha trifallax was blunt end-repaired with  
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
oligonucleotides were ligated to the blunt ends in high  
molar excess. Vector DNA was prepared from a derivative of  
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
derivative of plasmid R1. The vector was ligated with  
adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. Coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 101; DB 9; Length 518;  
Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60  
Db 410 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 351  
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101  
Db 350 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 310

RESULT 15  
LOCUS  
DEFINITION

CC817162 519 bp DNA linear GSS 17-JUL-2003  
100002J19 Oxytricha plasmid UUGC10 library Sterkiella  
histriomuscorum genomic clone UUGC100002J19 R, genomic survey  
sequence.  
CC817162  
CC817162.1 GI:32896449  
GSS.  
Sterkiella histriomuscorum (Oxytricha trifallax)  
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;  
Stichotrichida; Oxytrichidae; Sterkiella.  
1 (bases 1 to 519)  
Dunn,D., Doak,T., Herrick,G. and Weiss,R.  
Paired end reads from plasmid inserts of Oxytricha trifallax  
macronuclear chromosomes  
Unpublished (2003)  
Contact: Robert B. Weiss  
University of Utah Genome Center

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Plate: 0002 row: J column: 19  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 519.

FEATURES  
Location/Qualifiers  
1..519  
/organism="Sterkiella histriomuscorum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:94289"  
/clone="UUGC100002J19"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Oxytricha plasmid UUGC10 library"  
/note="Vector: PWD42nv; Purified macronuclear chromosomal  
DNA from Oxytricha trifallax was blunt end-repaired with  
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
oligonucleotides were ligated to the blunt ends in high  
molar excess. Vector DNA was prepared from a derivative of  
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
derivative of plasmid R1. The vector was ligated with  
adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. Coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 101; DB 9; Length 519;  
Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60  
Db 416 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 357  
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101  
Db 356 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 316

Search completed: July 14, 2005, 23:23:05  
Job time : 962.667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1\_gacggatcgggagatctccc.....ctgtccctgttgtgtgt 100

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_bts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3986	12 PCDNA32EO	X90639 Cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Expressio
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5082	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CVU89673	U89673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A441171	A441171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222266	AR222266 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CVU89672	U89672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1  
AR098190  
LOCUS AR098190 3853 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 5 from patent US 6074850.  
ACCESSION AR098190  
VERSION AR098190.1 GI:12807447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3853)  
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.  
TITLE Retinoblastoma fusion polypeptides  
JOURNAL Patent: US 6074850-A 5.13-JUN-2000;  
FEATURES Location/Qualifiers  
source 1..3853  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 100.0%; Score 100; DB 6; Length 3853;  
Best Local Similarity 100.0%; Pred. No. 9.4e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCATCTCAGTACATCTGCTCTGATG 60  
Db 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCATCTCAGTACATCTGCTCTGATG 60  
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTT 100  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTT 100  
RESULT 2  
AR207832  
LOCUS AR207832 3853 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 5 from patent US 6379927.  
ACCESSION AR207832  
VERSION AR207832.1 GI:21507688  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

```
Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 3
BD009729
LOCUS BD009729 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified

REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANUI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 209..862.
FEATURES source
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

PCDNA3ZEO
LOCUS PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995

REFERENCE 1 (bases 1 to 3986)
AUTHORS Peters,H., Hunthausen,T., Kroenke,M. and Marget,M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters,H.
TITLE Direct Submission
JOURNAL Michaelisstr.5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES Location/Qualifiers
source 1..3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pcDNA3ZEO"
misc_feature 1..2125
/notes="cloning vector (pcDNA3) (Invitrogen)"
misc_feature 889..994
/notes="multiple cloning site (MCS)"
misc_feature 2126..2796
/notes="cloning vector (PzeoSV) (Invitrogen)"
misc_feature 2797..3986
/notes="cloning vector (pcDNA3)"

ORIGIN
Query Match 100.0%; Score 100; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 5
AR098191
LOCUS AR098191 4026 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..4026
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
```

```

Query Match          100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
QY     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 8
AR098192
LOCUS      4249 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION    AR098192.1 GI:12807449
KEYWORDS
SOURCE
ORGANISM   Unknown.
           Unclassified.
REFERENCE  1 (bases 1 to 4249)
AUTHORS   Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE     Retinoblastoma fusion polypeptides
JOURNAL   Patent: US 6074850-A 33 13-JUN-2000;
FEATURES   Location/Qualifiers
            source          1. .4249
                        /organism="unknown"
                        /mol_type="unassigned DNA"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
QY     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9
AR207834
LOCUS      4249 bp      DNA      linear      PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION    AR207834.1 GI:21507691
KEYWORDS
SOURCE
ORGANISM   Unknown.
           Unclassified.
REFERENCE  1 (bases 1 to 4249)
AUTHORS   Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE     Retinoblastoma fusion proteins
JOURNAL   Patent: US 6379927-A 33 30-APR-2002;
FEATURES   Location/Qualifiers
            source          1. .4249
                        /organism="unknown"
                        /mol_type="unassigned DNA"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db      1  GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
QY     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db     61  CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249
FT /organism='Unidentified'.
FEATURES
source
1..4249
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES
source
1..4341
Location/Qualifiers
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'
```

```
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source
1..4341
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1..4597
Location/Qualifiers
/organism='synthetic construct'
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 14
AXI33940
LOCUS AXI33940 4840 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AXI33940
VERSION AXI33940.1 GI:14139881
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE Cell transfection
JOURNAL Patent: WO 0119853-A 1 22-MAR-2001;
THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES
source
1. 4840
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="This sequence is artificial and is based on well
established commercially available vectors that are cited
with their vendor within the patent application"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 15
BD238492
LOCUS BD238492 5053 bp DNA linear PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
of using the same.
ACCESSION BD238492
VERSION BD238492.1 GI:33048262
KEYWORDS JP 2002520000-A/18.
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
and Chesnut,R.W.
TITLE Expression vectors for stimulating an immune response and methods
of using the same
JOURNAL Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904, 15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
PC A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
/organism="Artificial Sequence".
FEATURES
source
1..5053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

Search completed: July 14, 2005, 14:03:32
Job time : 749.127 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright. (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_1\_100

Perfect score: 100

Sequence: 1 gacggatcggagatctccc.....ctgtccctcgtgtgtgtt 100

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002as:\*

7: Geneseq2002bs:\*

8: Geneseq2003as:\*

9: Geneseq2003bs:\*

10: Geneseq2003cs:\*

11: Geneseq2003ds:\*

12: Geneseq2004as:\*

13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	1506	12 ADMA1035	Adm1035 Fungus nu
2	100	100.0	1600	12 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADMA1037	Adm1037 Cytomegal
4	100	100.0	2241	12 ADMA1034	Adm1034 Human nuc
5	100	100.0	2294	12 ADMA1036	Adm1036 Cytomegal
6	100	100.0	3853	2 AAV40006	Aav40006 Plasmid p
7	100	100.0	4026	2 AAV40007	Aav40007 Plasmid p
8	100	100.0	4249	2 AAV63466	Aav63466 Plasmid p
9	100	100.0	4341	2 AAV62391	Aav62391 Vector pv
10	100	100.0	4341	2 AAS17704	Aas17704 Vector pv
11	100	100.0	4341	6 AEN83143	Abn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pEF2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA sequ
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rd

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 AD821866	Ad821866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AAZ89476	Aaz89476 Transgeni
33	100	100.0	5446	6 AAS18619	Aas18619 Renilla l
34	100	100.0	5446	6 ABL53540	AbL53540 Vector pc
35	100	100.0	5446	12 ADM36314	Adm36314 Plasmid p
36	100	100.0	5458	6 ABL58494	AbL58494 Recombina
37	100	100.0	5458	6 ABL58493	AbL58493 Recombina
38	100	100.0	5543	6 ABK88868	Abk88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ad83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	AbL58489 Recombina
42	100	100.0	5614	6 ABL58490	AbL58490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AA166195	Aa166195 Human FSH
45	100	100.0	5651	6 ABK40237	Abk40237 DNA encod

#### ALIGNMENTS

##### RESULT 1

ADMA1035

ID ADMA1035 standard; DNA; 1506 BP.

XX AC ADMA1035;

XX DT 17-JUN-2004 (first entry)

XX DE Fungus nucleotide sequence SEQ ID NO:3.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;  
growth; differentiation; drug development; vaccine development;  
tissue transplantation; human disease study; fungus; gene; ds.  
XX OS Unidentified.  
XX FN WO2004027029-A2.  
XX PD 01-APR-2004.  
XX PF 17-SEP-2003; 2003WO-US029251.  
XX PR 19-SEP-2002; 2002US-0411790P.  
XX PA (XIME-) XIMEREX INC.  
XX PI Beechornor WE, Sosa CE, Thompson SC;  
XX DR WPI; 2004-295402/27.  
XX PT Engrafting foreign replacement cells within a fetal non-human mammal,  
useful in producing chimeric mammals, comprises selectively destroying  
native cells in a tissue of a fetal non-human mammal host.  
XX PS Disclosure; SEQ ID NO 3; 48pp; English.  
XX CC The present invention describes a method for engrafting foreign  
replacement cells within a foetal non-human mammal, which comprises  
selectively destroying native cells in a tissue of a foetal non-human  
mammal host, where the number of maternal cells of the same tissue is not  
substantially reduced, and implanting foreign replacement cells in the  
tissue of the fetal non-human mammal host, where the foreign replacement  
cells replace destroyed cells of the tissue. The method is useful for



CC replacement cells within a foetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a foetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the foetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1782;  
Best Local Similarity 100.0%; Pred. No. 4.2e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

## RESULT 4

ADM41034  
ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

XX DT 17-JUN-2004 (first entry)

XX DE Human nucleotide sequence SEQ ID NO:2.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;  
XX KW growth; differentiation; drug development; vaccine development;  
XX KW tissue transplantation; human disease study; human; gene; ds.

XX OS Homo sapiens.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX FA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX PT Engrafting foreign replacement cells within a foetal non-human mammal,  
XX PT useful in producing chimeric mammals, comprises selectively destroying  
XX PT native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 2; 48pp; English.

XX CC The present invention describes a method for engrafting foreign  
XX CC replacement cells within a foetal non-human mammal, which comprises  
XX CC selectively destroying native cells in a tissue of a foetal non-human  
XX CC mammal host, where the number of maternal cells of the same tissue is not  
XX CC substantially reduced, and implanting foreign replacement cells in the  
XX CC tissue of the foetal non-human mammal host, where the foreign replacement  
XX CC cells replace destroyed cells of the tissue. The method is useful for  
XX CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;  
Best Local Similarity 100.0%; Pred. No. 4.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

## RESULT 5

ADM41036  
ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

XX DT 17-JUN-2004 (first entry)

XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;  
XX KW growth; differentiation; drug development; vaccine development;  
XX KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX OS Cytomegalovirus.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX FA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX PT Engrafting foreign replacement cells within a foetal non-human mammal,  
XX PT useful in producing chimeric mammals, comprises selectively destroying  
XX PT native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 4; 48pp; English.

XX CC The present invention describes a method for engrafting foreign  
XX CC replacement cells within a foetal non-human mammal, which comprises  
XX CC selectively destroying native cells in a tissue of a foetal non-human  
XX CC mammal host, where the number of maternal cells of the same tissue is not  
XX CC substantially reduced, and implanting foreign replacement cells in the  
XX CC tissue of the foetal non-human mammal host, where the foreign replacement  
XX CC cells replace destroyed cells of the tissue. The method is useful for  
XX CC facilitating growth and differentiation of foreign cells within a  
XX CC mammalian host, and for producing chimeric mammals that can be used to  
XX CC develop new drugs and vaccine, factors, drugs and tissues for  
XX CC transplantation, also useful to study human diseases. The present  
XX CC sequence represents a nucleotide sequence given in the Sequence Listing  
XX CC of the present invention but not mentioned further within the  
XX CC specification.

<p> <b>SQ</b> Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;  <b>Query Match</b> 100.0%; Score 100; DB 12; Length 2294;  <b>Best Local Similarity</b> 100.0%; Pred. No. 4.5e-26;  <b>Matches</b> 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;         </p> <p> <b>Qy</b> 1 GACGGATCGGGAGATCTCCCGATCCCGATCGCTTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60                 <b>Db</b> 1 GACGGATCGGGAGATCTCCCGATCCCGATCGCTTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60                 <b>Qy</b> 61 CCCCATAGTTAAGCCAGTATCTGCCCTCCCTGCTTGTGTGTT 100                 <b>Db</b> 61 CCCCATAGTTAAGCCAGTATCTGCCCTCCCTGCTTGTGTGTT 100                        </p> <p> <b>RESULT 6</b>  <b>AAV40006</b>  <b>ID</b> AAV40006 standard; DNA; 3853 BP.  <b>XX</b>  <b>AC</b> AAV40006;  <b>XX</b>  <b>DT</b> 27-AUG-2003 (revised)  <b>DT</b> 15-FEB-1999 (first entry)  <b>XX</b>  <b>DE</b> Plasmid pCTM.  <b>XX</b>  <b>KW</b> E2F; transcription factor; human; retinoblastoma protein RB;  <b>KW</b> bladder cancer; restenosis; angioplasty; diabetic retinopathy;  <b>KW</b> thyroid hyperplasia; Grave's disease; psoriasis;  <b>KW</b> benign prostatic hypertrophy; Li-Fraumeni syndrome;  <b>KW</b> peripheral vascular disease; therapy; plasmid pCTM; ss.  <b>XX</b>  <b>OS</b> Human cytomegalovirus.  <b>OS</b> mastadenovirus.  <b>OS</b> unidentified bacteriophage; T7.  <b>OS</b> unidentified bacteriophage; SP6.  <b>OS</b> Macaca mulatta; polyoma virus.  <b>OS</b> Bos taurus.  <b>OS</b> Chimeric.  <b>XX</b>  <b>PH</b> Key Location/Qualifiers  <b>FT</b> promoter 209..864            /*tag= a  <b>FT</b> /*note= "CMV promoter"  <b>FT</b> misc_feature 907..1131            /*tag= b  <b>FT</b> /*function= "tripartite leader sequence"  <b>FT</b> promoter 1132..1149            /*tag= c  <b>FT</b> /*note= "SP6 promoter"  <b>FT</b> misc_feature 1679..3853            /*tag= d  <b>FT</b> /*note= "pUC19 backbone H3 to AatII"  <b>FT</b> CDS complement (2857..3717)            /*tag= e  <b>FT</b> /*note= "AMP-ORF"  <b>XX</b>  <b>PN</b> W0981228-A1.  <b>XX</b>  <b>XX</b>  <b>PD</b> 22-MAY-1998.  <b>XX</b>  <b>XX</b>  <b>PF</b> 13-NOV-1997; 97WO-US021821.  <b>XX</b>  <b>XX</b>  <b>PR</b> 15-NOV-1996; 96US-00751517.  <b>PR</b> 14-FEB-1997; 97US-00801092.  <b>XX</b>  <b>XX</b>  <b>PA</b> (CANJ-) CANJI INC.  <b>XX</b>  <b>PI</b> Antelman D, Gregory RJ, Wills KN;  <b>XX</b>  <b>XX</b>  <b>DR</b> WPI; 1998-297858/26.  <b>XX</b>  <b>XX</b> New fusion polypeptide of, e.g. transcription factor - used to treat,         </p>	<p> <b>PT</b> e.g. hyper-proliferative disease such as cancer and restenosis.  <b>XX</b>  <b>PS</b> Example 1; Fig 4; 91pp; English.  <b>XX</b>  <b>CC</b> This is the nucleotide sequence of pCTM, a plasmid which contains a CMV promoter, a tripartite adenovirus leader, flanked by T7 and SP6 promoters, and a multiple cloning site with a bovine growth hormone polyA site and downstream SV40 polyA site. It has been used as a vector for the expression of fusion proteins of the invention that comprise retinoblastoma protein (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such fusion proteins, particularly expressed from gene therapy vectors, are used to treat hyperproliferative conditions, specifically cancer (particularly of the bladder) or restenosis. They are more effective in repressing transcription of the E2F promoter than RB alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct OS field.)  <b>XX</b>  <b>SQ</b> Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;         </p> <p> <b>Query Match</b> 100.0%; Score 100; DB 2; Length 3853;  <b>Best Local Similarity</b> 100.0%; Pred. No. 5.2e-26;  <b>Matches</b> 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0         </p> <p> <b>Qy</b> 1 GACGGATCGGGAGATCTCCCGATCCCGATCGCTTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60                 <b>Db</b> 1 GACGGATCGGGAGATCTCCCGATCCCGATCGCTTATGGTCGACTCTCAGTACAATCTGCTCTGATG 60                 <b>Qy</b> 61 CCGCATAGTTAAGCCAGTATCTGCCCTCCCTGCTTGTGTGTT 100                 <b>Db</b> 61 CCGCATAGTTAAGCCAGTATCTGCCCTCCCTGCTTGTGTGTT 100                        </p> <p> <b>RESULT 7</b>  <b>AAV40007</b>  <b>ID</b> AAV40007 standard; DNA; 4026 BP.  <b>XX</b>  <b>AC</b> AAV40007;  <b>XX</b>  <b>DT</b> 27-AUG-2003 (revised)  <b>DT</b> 15-FEB-1999 (first entry)  <b>XX</b>  <b>DE</b> Plasmid pCTMI.  <b>XX</b>  <b>KW</b> E2F; transcription factor; human; retinoblastoma protein RB;  <b>KW</b> bladder cancer; restenosis; angioplasty; diabetic retinopathy;  <b>KW</b> thyroid hyperplasia; Grave's disease; psoriasis;  <b>KW</b> benign prostatic hypertrophy; Li-Fraumeni syndrome;  <b>KW</b> peripheral vascular disease; therapy; plasmid pCTMI; ss.  <b>XX</b>  <b>OS</b> Human cytomegalovirus.  <b>OS</b> mastadenovirus.  <b>OS</b> unidentified bacteriophage; T7.  <b>OS</b> unidentified bacteriophage; SP6.  <b>OS</b> Macaca mulatta; polyoma virus.  <b>OS</b> Bos taurus.  <b>OS</b> Chimeric.  <b>XX</b>  <b>PH</b> Key Location/Qualifiers  <b>FT</b> promoter 209..864            /*tag= a  <b>FT</b> /*note= "CMV promoter"  <b>FT</b> misc_feature 907..1074            /*tag= b  <b>FT</b> /*function= "tripartite leader sequence"  <b>FT</b> intron 1075..1253            /*tag= c  <b>FT</b> /*note= "hybrid SV40 late intron"  <b>FT</b> promoter 1305..1322            /*tag= d  <b>FT</b> /*note= "SP6 promoter"  <b>FT</b> misc_feature 1851..4026            /*tag= e  <b>FT</b> /*note= "pUC19 backbone H3 to AatII"         </p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```

FT CDS complement(3032..3890)
FT /*tag= f
FT /*note= "AMP-ORF"
XX
XX
XX PN W09821228-A1.
XX PD 22-MAY-1998.
XX
XX PF 13-NOV-1997; 97WO-US021821.
XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX PA (CANJ-) CANJI INC.
XX
XX PI Antelman D, Gregory RJ, Wills KN;
XX
XX DR WPI; 1998-297858/26.
XX
XX PT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX PS Example 1; Fig 6; 91pp; English.
XX
XX CC This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX CC from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX CC subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX CC vector. Plasmid pCTMI has been used as a vector for the expression of
XX CC fusion proteins of the invention that comprise retinoblastoma protein
XX CC (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX CC fusion proteins, particularly expressed from gene therapy vectors, are
XX CC used to treat hyperproliferative conditions, specifically cancer
XX CC (particularly of the bladder) or restenosis. They are more effective in
XX CC repressing transcription of the E2F promoter than RB alone and cause cell
XX CC -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX CC OS field.)
XX
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4026;
Best Local Similarity 100.0%; Pred. No. 5.3e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGATG 60
Oy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 8
AAV63466
ID AAV63466 standard; DNA; 4249 BP.
XX
XX AC AAV63466;
XX
XX DT 27-AUG-2003 (revised)
XX DT 15-FEB-1999 (first entry)
XX
XX DE Plasmid pCTMIE.
XX
XX KW E2F; transcription factor; human; retinoblastoma protein RB;
XX KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX KW thyroid hyperplasia; Grave's disease; psoriasis;
XX KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX KW peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX
XX OS Human cytomegalovirus.
XX OS mastadenovirus.
XX OS unidentified bacteriophage; T7.
XX OS unidentified bacteriophage; SP6.

```

---

```

OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
XX FH Key
XX FT Promoter
XX FT 209..864
XX FT /*tag= a
XX FT /*note= "CMV promoter"
XX FT 907..1074
XX FT /*tag= b
XX FT /function= "tripartite leader sequence"
XX FT 1081..1145
XX FT /*tag= c
XX FT /*note= "hybrid SV40 late intron"
XX FT 1164..1366
XX FT /*tag= d
XX FT /*note= "early mRNA"
XX FT 1261..1332
XX FT /*tag= e
XX FT /*note= "72 bp tandem repeat enhancer"
XX FT 1333..1404
XX FT /*tag= f
XX FT /*note= "72 bp tandem repeat enhancer"
XX FT 1366
XX FT /*tag= g
XX FT /*note= "T antigen binding site"
XX FT 1372..1478
XX FT /*tag= h
XX FT /*note= "hybrid SV40 late intron"
XX FT 1530..1545
XX FT /*tag= i
XX FT /*note= "SP6 promoter"
XX FT 2075..4249
XX FT /*tag= j
XX FT /*note= "pUC19 backbone H3 to AatII"
XX FT complement(3255..4113)
XX FT /*tag= k
XX FT /*note= "AMP-ORF"
XX
XX W09821228-A1.
XX
XX PD 22-MAY-1998.
XX
XX PF 13-NOV-1997; 97WO-US021821.
XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX PA (CANJ-) CANJI INC.
XX
XX PI Antelman D, Gregory RJ, Wills KN;
XX
XX DR WPI; 1998-297858/26.
XX
XX PT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX PS Example 1; Fig 8; 91pp; English.
XX
XX CC This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX CC by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX CC amplified product with BglII and inserting into BamHI-digested plasmid
XX CC pCMTI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX CC expression of fusion proteins of the invention that comprise
XX CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX CC (see AAW62464). Such fusion proteins, particularly expressed from gene
XX CC therapy vectors, are used to treat hyperproliferative conditions,
XX CC specifically cancer (particularly of the bladder) or restenosis. They are
XX CC more effective in repressing transcription of the E2F promoter than RB
XX CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX CC AUG-2003 to correct OS field.)
XX
XX SQ Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;

```



SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB |||||  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

RESULT 11  
ABN83143  
ID ABN83143 standard; DNA; 4341 BP.  
XX  
AC ABN83143;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Plasmid pVAC1 complete sequence.  
XX  
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;  
KW cancer; B cell malignancy; ds.  
XX  
OS Synthetic.  
XX  
PN WO200240513-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-GB005142.  
XX  
PR 20-NOV-2000; 2000GB-00028319.  
XX  
PS (CANC-) CANCER RES VENTURES LTD.  
XX  
PI Savelyeva N, Stevenson F;  
XX  
PT WPI; 2002-500202/53.  
XX  
PS Nucleic acid construct for delivery into living cells as a vaccine,  
PT useful for treating e.g. cancer, directs the expression of a fusion  
PT protein comprising an antigen and an adjuvant sequence derived from a  
PT plant viral coat protein.  
XX  
PS Example 3; Fig 7; 84pp; English.  
XX  
CC The invention relates to a novel nucleic acid construct for inducing an  
CC immune response in vivo to an antigen, capable of directing the  
CC expression of a fusion protein that comprises an antigen and an adjuvant  
CC sequence derived from a plant viral coat protein. The construct of the  
CC invention has cytostatic and virucide activity. The nucleic acid  
CC construct is useful for inducing an immune response in a patient, for  
CC vaccinating a patient against an infectious disease caused by an antigen  
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a  
CC patient with a predisposition to cancer and for treating a patient having  
CC a B cell malignancy, where the construct is encapsulated, and optionally,  
CC a second nucleic acid sequence encoding a further immunomodulatory  
CC polypeptide is administered to the patient. The construct is also useful  
CC in medical treatment, and in the preparation of a vaccine for treating or  
CC preventing a disease state associated with the antigen. The sequence  
CC shows the complete sequence of vector pVAC1  
XX  
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

RESULT 12  
AAF24901  
ID AAF24901 standard; DNA; 4597 BP.  
XX  
AC AAF24901;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.  
XX  
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;  
KW myocardial ischemia; cardiac angiogenesis; haemophilia;  
KW vascular endothelial growth factor; VEGF; ss.  
XX  
OS Synthetic.  
XX  
PN WO200078358-A2.  
XX  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000WO-US016837.  
XX  
PR 18-JUN-1999; 99US-0140260P.  
XX  
PS (COLL-) COLLABORATIVE GROUP LTD.  
XX  
PI Chen W;  
XX  
PT WPI; 2001-071363/08.  
XX  
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial  
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic  
PT acids crosslinked to nucleic acids.  
XX  
PS Example 1; Page 36-38; 38pp; English.  
XX  
CC The specification describes a microsphere comprising dihydrazide  
CC derivatized hyaluronic acid crosslinked to a nucleic acid (NA). The  
CC microspheres cause reduced inflammatory responses, and have increased  
CC safety and biodegradability. The microspheres are useful for transfecting  
CC a cell of a subject and for treating a subject having myocardial  
CC ischemia, by increasing cardiac angiogenesis. They are also useful for  
CC treating haemophilia. The present sequence represents the plasmid  
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is  
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a  
CC vascular endothelial growth factor (VEGF)  
XX  
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 4; Length 4597;  
Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB |||||

RESULT 13  
AAD39652







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_1\_100

Perfect score: 100

Sequence: 1 gacggatcggagatctccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	60.0	602	8	B67169 CpG0047A Cp
2	55.6	55.6	694	8	B2052929 jnr13g03.
3	55.6	55.6	696	8	B2050328 jnr42c12.
4	55.6	55.6	717	8	B2054067 jnr38b09.
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK113937 CK113937
9	53.4	53.4	766	7	CK120360 CK120360
10	53.4	53.4	788	7	CK117844 CK117844
11	53.4	53.4	898	9	CL141237 ISB1-118J
12	53.4	53.4	899	9	CL140877 ISB1-118B
13	53.4	53.4	1009	9	CL123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbe0049M
15	53	53.0	675	8	B2051815 jnr57d03.
16	53	53.0	679	8	B2052857 jnr13g03.
17	53	53.0	700	8	B2050646 jnr66f08.
18	53	53.0	701	8	B2052015 jnr56b03.
19	53	53.0	708	8	B2054793 jnr33g03.
20	53	53.0	709	8	B2053587 jnr98d01.
21	53	53.0	712	8	B2054005 jnr38b09.
22	52.8	52.8	451	8	AQ863966 nbe00022E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	B2049113 jnr21d02.

25	52.4	52.4	708	8	BZ050047
26	51.6	51.6	328	9	CC819886
27	51.6	51.6	351	9	CC818492
28	51.6	51.6	358	9	CC817661
29	51.6	51.6	364	9	CC817805
30	51.6	51.6	364	9	CC818511
31	51.6	51.6	364	9	CC818574
32	51.6	51.6	364	9	CC819049
33	51.6	51.6	369	9	CC817069
34	51.6	51.6	374	9	CC817074
35	51.6	51.6	374	9	CC820036
36	51.6	51.6	395	9	CC817652
37	51.6	51.6	403	9	CC817682
38	51.6	51.6	403	9	CC817837
39	51.6	51.6	414	9	CC819240
40	51.6	51.6	419	9	CC818384
41	51.6	51.6	420	9	CC817834
42	51.6	51.6	426	9	CC817720
43	51.6	51.6	437	9	CC819820
44	51.6	51.6	441	9	CC818421
45	51.6	51.6	443	9	CC817769

## ALIGNMENTS

RESULT 1  
B67169  
LOCUS CP60047A CptIOWAgDNA2 602 bp DNA linear GSS 12-MAY-2000  
DEFINITION Cryptosporidium parvum genomic, genomic survey  
sequence.  
ACCESSION B67169  
VERSION B67169.1 GI:2642750  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum  
ORGANISM Cryptosporidium parvum  
Cryptosporidium parvum  
Cryptosporidiidae; Cryptosporidium.  
REFERENCE  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1); 1-32 (2000)  
MEDLINE 20183851  
PUBMED 10717299  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malariaditsa.ucsf.edu  
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.  
Seq primer: T7  
Class: shotgun  
High quality sequence stop: 602.

## FEATURES

source  
Location/Qualifiers  
1..602  
/organism="Cryptosporidium parvum"  
/mol\_type="genomic DNA"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/lab\_host="E. coli XL2 Blue MRF"  
/clone\_lib="CptIOWAgDNA2"  
/notes="vector: PCR-script Amp SK+; Site\_1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested pCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

## ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;  
Best Local Similarity 100.0%; Pred. No. 2.4e-10;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAACTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100  
|||||  
Db 1 CAGTACAACTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 60  
|||||

RESULT 2  
BZ052929/c  
LOCUS  
DEFINITION jnr13903.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
sequence.

ACCESSION  
BZ052929  
VERSION  
BZ052929.1 GI:23654922  
KEYWORDS  
GSS.

SOURCE  
Brassica oleracea  
ORGANISM  
Brassica oleracea

REFERENCE  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr13 row: 9 column: 03

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

## FEATURES

source  
Location/Qualifiers  
1..694  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"  
/note="Vector: pOTw13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using Brassica oleracea T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;  
Best Local Similarity 77.9%; Pred. No. 9e-09;  
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGGTGCATCTCAGTACAATCTGCTCTGATGCC 62  
|||||  
Db 324 CGGATCGATAGGTCCTCGGATAGTTATGGTGCATCTCAGTACAATCTGCTCTGATGCC 265  
|||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
|||||

Db 264 GCATAGTTAAGCCAGCCCGCACACCC 239

## RESULT 3

BZ050328

LOCUS

DEFINITION

BZ050328 696 bp DNA linear GSS 09-OCT-2002

jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey

sequence.

ACCESSION

BZ050328

VERSION

BZ050328.1 GI:23649718

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: c column: 12

Seq primer: -21UPPOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

Location/Qualifiers

1..696

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/db\_xref="taxon:3712"

/clone\_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from

flowering buds. DNA was purified from a crude nuclear

prep using Brassica oleracea T01000DH3 buds provided by

Thomas Osborn at the University of Wisconsin. Genomic

DNA was provided by Pablo Rabinowicz (CSHL) and the

shotgun library prepared at Washington University Genome

Sequencing Center."

Query Match 55.6%; Score 55.6; DB 8; Length 696;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGGTGCATCTCAGTACAATCTGCTCTGATGCC 62

|||||

Db 45 CGGATCGATAGGTCCTCGGATAGTTATGGTGCATCTCAGTACAATCTGCTCTGATGCC 104

|||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

|||||

Db 105 GCATAGTTAAGCCAGCCCGCACACCC 130

|||||

RESULT 4

BZ054067/c

LOCUS

DEFINITION

jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

sequence.

ACCESSION

BZ054067

VERSION

BZ054067.1 GI:23657216

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
Whole genome shotgun reads from *Brassica oleracea*  
Unpublished (2002)  
Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Plate: jnr38 row: b column: 09  
Seq primer: -28RPpOT reverse  
Class: shotgun  
High quality sequence start: 87  
High quality sequence stop: 543.  
Location/Qualifiers

# FEATURES

1. .717  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea001"  
/note="Vector: pOTw13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using *Brassica oleracea* T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;  
Best Local Similarity 77.9%; Pred. No. 9.1e-09;  
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
Qy 3 CGGATCGGGAGATCTCCGATGCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 62  
Db CGGATCGATAGTCTCCGCTGACTAGTTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 189  
Qy 63 GCATAGTTAAGCCAGTATCTGCTGCC 88  
Db 188 GCATAGTTAAGCCAGCCGACACCC 163

## RESULT 5

AW409112 348 bp mRNA linear EST 31-DEC-2000  
LOCUS sal10h5 Salivary Gland Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION AW409112  
ACCESSION AW409112  
VERSION AW409112.1 GI:11999687  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 348)  
Bergheim, A., Ogilvie, E., Arndt, S., Napier, H., Taylor, M., Lovett, M.,  
Simmons, A., Hide, W., and Ramsay, M.  
A high density transcript map between markers D8S550 and D8S1759 on  
8p22-23, using cDNA direct selection  
Unpublished (2000)  
Contact: Ramsey M  
Department of Human Genetics  
South African Institute For Medical Research  
P.O.Box 1038, Johannesburg, Gauteng, 2000, South Africa  
Fax: 2711 489 9226  
Email: micheler@mail.saimr.wits.ac.za.

## Location/Qualifiers

1. .348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Salivary Gland"  
/clone\_lib="Salivary Gland Library"  
/note="Vector: pAMP10"

## FEATURES

### source

AL714571/c  
LOCUS AL714571  
DEFINITION AL714571 Danio rerio embryonic inner ear subtracted cDNA Danio  
rerio cDNA clone BNOAA007ZC02 5', mRNA sequence.  
ACCESSION AL714571  
VERSION AL714571.1 GI:20179174  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## ORIGIN

Query Match 53.4%; Score 53.4; DB 1; Length 343;  
Best Local Similarity 84.5%; Pred. No. 4.8e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTTAAGCCA 76  
Db 280 TTCACACCGCATATGGTGCCTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTTAAGCCA 221  
Qy 77 GTATCTGCTCC 87  
Db 220 GTATACACTCC 210

## RESULT 7

AL714571/c  
LOCUS AL714571  
DEFINITION AL714571 Danio rerio embryonic inner ear subtracted cDNA Danio  
rerio cDNA clone BNOAA007ZC02 5', mRNA sequence.  
ACCESSION AL714571  
VERSION AL714571.1 GI:20179174  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 53.6%; Score 53.6; DB 2; Length 348;  
Best Local Similarity 80.5%; Pred. No. 4.1e-08;  
Matches 62; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
Qy 11 GAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTT 70  
Db 65 GCGGTATACACCGCATATGGTGCCTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTT 124  
Qy 71 AAGCCAGTATCTGCTCC 87  
Db 125 AAGCCAGTATACACTCC 141

## RESULT 6

AL715724/c  
LOCUS AL715724  
DEFINITION AL715724 Danio rerio embryonic inner ear subtracted cDNA Danio  
rerio cDNA clone BNOAA018ZF12 5', mRNA sequence.  
ACCESSION AL715724  
VERSION AL715724.1 GI:20180327  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

## REFERENCE

### AUTHORS

1 (bases 1 to 343)  
Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,  
Hardelin, J.P., Weissenbach, J. and Petit, C.

### TITLE

A subtracted cDNA library from the zebrafish (Danio rerio)  
embryonic inner ear

### JOURNAL

Unpublished (2002)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

### FEATURES

#### source

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone\_lib="BNOAA018ZF12"  
/tissue\_type="inner ear"  
/dev\_stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cDNA"  
/note="subtracted cDNA library"

## ORIGIN

Query Match 53.4%; Score 53.4; DB 1; Length 343;  
Best Local Similarity 84.5%; Pred. No. 4.8e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTTAAGCCA 76  
Db 280 TTCACACCGCATATGGTGCCTCTCAGTACAAATCTGCTGTGATCCGCGCATAGTTAAGCCA 221  
Qy 77 GTATCTGCTCC 87  
Db 220 GTATACACTCC 210

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS 1 (bases 1 to 345)
TITLE A substracted cDNA library from the zebrafish (Danio rerio)
JOURNAL embryonic inner ear
COMMENT Unpublished (2002)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source
1..345
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA0072C02"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear substracted
cDNA"
/note="substracted cDNA library"
ORIGIN
Query Match 53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76
Db 280 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 221
QY 77 GTATCTGCTCC 87
Db 220 GTATACACTCC 210
RESULT 8
CK119397/c
LOCUS CK119397 761 bp mRNA linear EST 01-JUN-2004
DEFINITION 212009, pl AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011J009212
5-PRIME, mRNA sequence.
ACCESSION CK119397
VERSION CK119397.1 GI:47829713
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
TITLE Generation of a cDNA expression library from Arabidopsis
inflowescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73 D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 761 Std Error: 0.00
Plate: 212 row: 0 column: 9
Seq primer: POE65.
Location/Qualifiers
1..761
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:954234"
/db_xref="taxon:3702"
FEATURES
source
1..761
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:954234"
/db_xref="taxon:3702"
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS 1 (bases 1 to 345)
TITLE A substracted cDNA library from the zebrafish (Danio rerio)
JOURNAL embryonic inner ear
COMMENT Unpublished (2002)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source
1..345
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA0072C02"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear substracted
cDNA"
/note="substracted cDNA library"
ORIGIN
Query Match 53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76
Db 280 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 221
QY 77 GTATCTGCTCC 87
Db 220 GTATACACTCC 210
RESULT 9
CK120360/c
LOCUS CK120360 766 bp mRNA linear EST 01-JUN-2004
DEFINITION 207J04, pl AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011J04207
5-PRIME, mRNA sequence.
ACCESSION CK120360
VERSION CK120360.1 GI:47830676
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
TITLE Generation of a cDNA expression library from Arabidopsis
inflowescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73 D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 766 Std Error: 0.00
Plate: 207 row: J column: 4
Seq primer: POE65.
Location/Qualifiers
1..766
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:953059"
/db_xref="taxon:3702"
/clone="MPMPGP2011J04207"
/tissue_type="inflowescence meristem"
/dev_stage="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="Vector: POE-30NAST-attB (AY386205); Site_1: SalI;

```

Site 2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearranged into the sublibrary (plate numbers begin with 201) containing 5,000 putative expression clones. Average insert size is 1 kb. Note: The rearranged sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 766;  
Best Local Similarity 84.5%; Pred. No. 5.6e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 76  
Db 679 TTACACCGCATATGGTGCACCTCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 620  
Qy 77 GTATCTGCTCC 87  
Db 619 GTATACACTCC 609

## RESULT 10

CL117844/c  
LOCUS CL117844 788 bp mRNA linear EST 01-JUN-2004  
DEFINITION 209P08.p1 AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011P08209  
5-PRIME, mRNA sequence.

ACCESSION CL117844  
VERSION CL117844.1 GI:47828160  
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 788)

REFERENCE Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.  
AUTHORS Generation of a cDNA expression library from Arabidopsis inflorescence meristem  
TITLE Unpublished (2003)

JOURNAL Contact: Birgit Kersten  
COMMENT Plant Protein Chip Group, Department Lehrach  
Max-Planck-Institute for Molecular Genetics  
Innestr. 73, D-14195 Berlin, Germany

Tel: +49(0)30/84131648  
Fax: +49(0)30/84131128  
Email: Kersten@molgen.mpg.de  
Insert Length: 788 Std Error: 0.00  
Plate: 209 row: P column: 8  
Seq primer: pOE65.

## FEATURES

Location/Qualifiers  
1..788  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="GABI:953578"  
/db\_xref="taxon:3702"  
/clone="MPMPGP2011P08209"  
/tissue\_type="inflorescence meristem"  
/dev\_stage="about one week after bolting"  
/lab\_host="E. coli SCS-1/pSE111"  
/clone\_lib="AtM1"

/note="Vector: pOE-3ONAST-attB (AY386205); Site 1: SalI; Site 2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearranged into the sublibrary (plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The rearranged sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

## ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 788;  
Best Local Similarity 84.5%; Pred. No. 5.7e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 76  
Db 514 TTACACCGCATATGGTGCACCTCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 455  
Qy 77 GTATCTGCTCC 87  
Db 454 GTATACACTCC 444

## RESULT 11

CL141237/c  
LOCUS CL141237 898 bp DNA linear GSS 05-JAN-2004  
DEFINITION ISB1-118J17.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118J17,  
genomic survey sequence.

ACCESSION CL141237  
VERSION CL141237.1 GI:40634872  
KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 898)

AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE A physical map of the xenopus tropicalis genome  
JOURNAL Unpublished (2003)

COMMENT Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Insert Length: 75000 Std Error: 0.00  
Seq primer: T7 TAATACGACTCACTATAGG  
Class: BAC ends

High quality sequence start: 4  
High quality sequence stop: 742.

FEATURES  
Location/Qualifiers

1..898  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8364"  
/clone="ISB1-118J17"  
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"

## ORIGIN

Query Match 53.4%; Score 53.4; DB 9; Length 898;  
Best Local Similarity 84.5%; Pred. No. 5.8e-08;  
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 76  
Db 195 TTACACCGCATATGGTGCACCTCTCAGTACAACTCTCTGATGCCGATAGTTAAGCCA 136  
Qy 77 GTATCTGCTCC 87  
Db 135 GTATACACTCC 125

## RESULT 12

CL140877/c

```

LOCUS      CL140877              899 bp    DNA          linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
            genomic survey sequence.
ACCESSION  CL140877
VERSION    CL140877.1 GI:40634512
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 899)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   source
            Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /clone_lib="ISB1"
                /notes="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
            Query Match      53.4%; Score 53.4; DB 9; Length 899;
            Best Local Similarity 84.5%; Pred. No. 5.8e-08;
            Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY      17 TCCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
            |||||
DB      195 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 136
QY      77 GTATCTGCTCC 87
            |||||
DB      135 GTATACACTCC 125

RESULT 13
CL123953/c
LOCUS      CL123953              1009 bp    DNA          linear    GSS 05-JAN-2004
DEFINITION ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
            genomic survey sequence.
ACCESSION  CL123953
VERSION    CL123953.1 GI:40617588
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1009)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   source
            Location/Qualifiers
                1..1009
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-84J15"
                /clone_lib="ISB1"
                /notes="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
            Query Match      53.4%; Score 53.4; DB 9; Length 1009;
            Best Local Similarity 84.5%; Pred. No. 5.9e-08;
            Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY      17 TCCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
            |||||
DB      252 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 193
QY      77 GTATCTGCTCC 87
            |||||
DB      192 GTATACACTCC 182

RESULT 14
LOCUS      AQ914559              814 bp    DNA          linear    GSS 02-DEC-1999
DEFINITION nhe0049W21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
            cultivar-group) genomic clone nhe0049W21r, genomic survey
            sequence.
ACCESSION  AQ914559
VERSION    AQ914559.1 GI:6511075
KEYWORDS   GSS.
SOURCE     Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1 (bases 1 to 814)
REFERENCE  1 Wing, R. A. and Dean, R. A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nhe0049W21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /notes="Vector: pBACindigo; Site_1: EcorI; Site_2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)

6468.225 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_3944\_4044

Perfect score: 101

Sequence: 1 aagcggcgtcgagtctag.....gtttgccccctcccccggtgcc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	6050	6	AX195205 Sequence
2	101	100.0	6567	6	AX128350 Sequence
3	101	100.0	6623	6	AX128345 Sequence
4	101	100.0	6639	6	AX128351 Sequence
5	101	100.0	6649	6	AX180726 Sequence
6	101	100.0	6695	6	AX128347 Sequence
7	101	100.0	6695	6	AX128353 Sequence
8	101	100.0	6695	6	AX128354 Sequence
9	101	100.0	6746	6	AX128344 Sequence
10	101	100.0	6818	6	AX128346 Sequence
11	101	100.0	6828	6	AX128340 Sequence
12	101	100.0	6833	6	AX128349 Sequence
13	101	100.0	6900	6	AX128341 Sequence
14	101	100.0	6956	6	AX128348 Sequence
15	101	100.0	7038	6	AX128342 Sequence
16	101	100.0	7960	6	BD268233 Adenoviru
17	101	100.0	7989	6	BD268236 Adenoviru
18	100	99.0	7231	6	BD268252 Adenoviru
19	99	98.0	280	6	CQ788637 Sequence

20	99	98.0	4291	12	AF302189
21	99	98.0	5082	6	A91754
22	99	98.0	5082	6	BD085110
23	99	98.0	5432	6	BD234590
24	99	98.0	5432	6	AX026821
25	99	98.0	5650	6	AX226281
26	99	98.0	5731	6	AX202478
27	99	98.0	5771	12	AF060226
28	99	98.0	6082	6	AR278592
29	99	98.0	6082	6	AR367288
30	99	98.0	6082	6	AR400324
31	99	98.0	6082	6	AR405591
32	99	98.0	6082	6	AR563971
33	99	98.0	6082	6	AX141045
34	99	98.0	6082	6	AX200905
35	99	98.0	6082	6	AX267561
36	99	98.0	6365	6	AX511381
37	99	98.0	6759	6	CQ788635
38	99	98.0	6759	6	CQ788642
39	99	98.0	6801	6	AX128352
40	99	98.0	6801	6	AX128355
41	99	98.0	7108	6	E36262
42	99	98.0	7108	6	AX001326
43	99	98.0	7427	6	CQ768745
44	99	98.0	7493	6	CQ768742
45	99	98.0	8578	6	AR409005

## ALIGNMENTS

RESULT 1  
LOCUS AX195205 6050 bp DNA linear PAT 28-AUG-2001  
DEFINITION Sequence 10 from Patent WO0151626.  
ACCESSION AX195205  
VERSION AX195205.1 GI:15385768  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Lu, X., Sun, L. and Zhang, Y.  
TITLE Novel plasmid dna vectors  
JOURNAL Patent: WO 0151626-A 10 19-JUL-2001;  
ELIM BIOPHARMACEUTICALS, INC. (US)  
FEATURES  
source  
1. 6050  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="commercial plasmid"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6050;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCGCGGTCGAGTCTAGAGGCGCGGTTTAAACCCGCTGATCAGCTCGCTGCT 60  
Db 3828 AAGCGCGGTCGAGTCTAGAGGCGCGGTTTAAACCCGCTGATCAGCTCGCTGCT 3887  
Qy 61 TCTAGTTCAGCAGCATCTGTTGTTGCGCTCCCGCGTCC 101  
Db 3888 TCTAGTTCAGCAGCATCTGTTGTTGCGCTCCCGCGTCC 3928  
RESULT 2  
AX128350  
LOCUS AX128350 6567 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 11 from Patent WO0130843.  
ACCESSION AX128350  
VERSION AX128350.1 GI:14134863

KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 11 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES  
Location/Qualifiers  
source  
1..6567  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct E2CLBDAS"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 6567;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2171 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2230  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101  
|||||  
Db 2231 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2271  
|||||

RESULT 3  
AX128345  
LOCUS AX128345 6623 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 6 from Patent WO0130843.  
ACCESSION AX128345  
VERSION AX128345.1 GI:14134858  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 6 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES  
Location/Qualifiers  
source  
1..6623  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct C7LBDAS"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 6623;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2227 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2286  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101  
|||||  
Db 2287 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2327  
|||||

RESULT 4  
AX128351  
LOCUS AX128351 6639 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 12 from Patent WO0130843.  
ACCESSION AX128351  
VERSION AX128351.1 GI:14134864  
KEYWORDS

SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 12 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES  
Location/Qualifiers  
source  
1..6639  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct E2CLBDBS"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 6639;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2243 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2302  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101  
|||||  
Db 2303 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2343  
|||||

RESULT 5  
AX180726  
LOCUS AX180726 6649 bp DNA linear PAT 06-AUG-2001  
DEFINITION Sequence 6 from Patent WO0146694.  
ACCESSION AX180726  
VERSION AX180726.1 GI:15132581  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Joly,E.  
TITLE A bioluminescence resonance energy transfer (bret) fusion molecule  
and method of use  
JOURNAL Patent: WO 0146694-A 6 28-JUN-2001;  
Biosignal Packard Inc. (CA)

FEATURES  
Location/Qualifiers  
source  
1..6649  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="DNA sequence for Rluc-PKA-EYFP construct"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 6649;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2611 AAGCGCGCTCGAGTCTAGAGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2670  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101  
|||||  
Db 2671 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2711  
|||||

RESULT 6  
AX128347  
LOCUS AX128347 6695 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 8 from Patent WO0130843.  
ACCESSION AX128347  
VERSION AX128347.1 GI:14134860  
KEYWORDS

SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 8 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)  
FEATURES Location/Qualifiers  
source  
1..6695  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct C7LBD8S"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2358  
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 101  
|||||  
Db 2359 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 2399  
|||||

RESULT 7  
AX128353  
LOCUS AX128353 6695 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 14 from Patent WO0130843.  
ACCESSION AX128353  
VERSION AX128353.1 GI:14134866  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 14 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)  
FEATURES Location/Qualifiers  
source  
1..6695  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct LBDBSG400V"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2358  
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 101  
|||||  
Db 2359 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 2399  
|||||

RESULT 8  
AX128354  
LOCUS AX128354 6695 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 15 from Patent WO0130843.  
ACCESSION AX128354  
VERSION AX128354.1 GI:14134867  
KEYWORDS  
SOURCE synthetic construct

ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 15 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)  
FEATURES Location/Qualifiers  
source  
1..6695  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct LBDBSG521R"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2358  
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 101  
|||||  
Db 2359 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 2399  
|||||

RESULT 9  
AX128344  
LOCUS AX128344 6746 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 5 from Patent WO0130843.  
ACCESSION AX128344  
VERSION AX128344.1 GI:14134857  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 5 03-MAY-2001;  
Novartis AG (CH); The Scripps Research Institute (US)  
FEATURES Location/Qualifiers  
source  
1..6746  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct C7LBDAL"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6746;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2350 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2409  
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 101  
|||||  
Db 2410 TCTAGTTGCCAGCCATCTGTTTGGCCCTCCCGCGTCC 2450  
|||||

RESULT 10  
AX128346  
LOCUS AX128346 6818 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 7 from Patent WO0130843.  
ACCESSION AX128346  
VERSION AX128346.1 GI:14134859  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct

```
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 7 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6818
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDBL"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6818;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2422 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2481
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2482 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2522

RESULT 11
AX128340
LOCUS AX128340 6828 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130843.
ACCESSION AX128340
VERSION AX128340.1 GI:14134853
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 1 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6828
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct 2C7LBDBS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6828;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2432 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2491
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2492 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2532

RESULT 12
AX128349
LOCUS AX128349 6833 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 10 from Patent WO0130843.
ACCESSION AX128349
VERSION AX128349.1 GI:14134862
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 10 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6833
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDBS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6833;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2437 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2496
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2497 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2537

RESULT 13
AX128341
LOCUS AX128341 6900 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 2 from Patent WO0130843.
ACCESSION AX128341
VERSION AX128341.1 GI:14134854
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 2 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6900
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct 2C7LBDBS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6900;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2504 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2563
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2564 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2604

RESULT 14
AX128348
LOCUS AX128348 6956 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 9 from Patent WO0130843.
ACCESSION AX128348
VERSION AX128348.1 GI:14134861
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
other sequences; artificial sequences.
REFERENCE 1
```

AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 9 03-MAY-2001;  
Novartis AG (CH) ; The Scripps Research Institute (US)  
FEATURES  
source 1..6956  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct C7LBDCL"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 6956;  
Best Local Similarity 100.0%; Pred. No. 3.2e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCGGCGCTCGAGTCTAGAGGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2560 AAGCGGCGCTCGAGTCTAGAGGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2619  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTCC 101  
|||||  
Db 2620 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTCC 2660  
|||||  
RESULT 15  
AX128342  
LOCUS AX128342 7038 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 3 from Patent WO0130843.  
ACCESSION AX128342  
VERSION AX128342.1 GI:14134855  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
other sequences; artificial sequences.  
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.  
TITLE Ligand activated transcriptional regulator proteins  
JOURNAL Patent: WO 0130843-A 3 03-MAY-2001;  
Novartis AG (CH) ; The Scripps Research Institute (US)  
FEATURES  
source 1..7038  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Construct 2C7LBDCL"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 7038;  
Best Local Similarity 100.0%; Pred. No. 3.2e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCGGCGCTCGAGTCTAGAGGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60  
|||||  
Db 2642 AAGCGGCGCTCGAGTCTAGAGGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2701  
|||||  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTCC 101  
|||||  
Db 2702 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTCC 2742  
|||||

Search completed: July 14, 2005, 14:03:32  
Job time : 756.618 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_3944\_4044

Perfect score: 101

Sequence: 1 aagcgccgtcagctag.....gtttgccccccccctggcc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2017	8 ADA05887	Ado42506 Human NOV
2	101	100.0	2017	12 ADNG3050	Adn63050 Human NOV
3	101	100.0	2017	12 ADO42504	Ado42504 Human NOV
4	101	100.0	2017	12 ADO42500	Ado42500 Human NOV
5	101	100.0	2022	12 ADO42506	Ado42506 Human NOV
6	101	100.0	6050	5 AAD10237	Aad10237 Commercia
7	101	100.0	6567	4 AAD06054	Aad06054 Plasmid E
8	101	100.0	6623	4 AAD06049	Aad06049 Plasmid C
9	101	100.0	6639	4 AAD06055	Aad06055 Plasmid E
10	101	100.0	6695	4 AAD06051	Aad06051 Plasmid C
11	101	100.0	6695	4 AAD06057	Aad06057 Plasmid C
12	101	100.0	6695	4 AAD06058	Aad06058 Plasmid C
13	101	100.0	6746	4 AAD06048	Aad06048 Plasmid C
14	101	100.0	6818	4 AAD06050	Aad06050 Plasmid C
15	101	100.0	6828	4 AAD06044	Aad06044 Plasmid C
16	101	100.0	6833	4 AAD06053	Aad06053 Plasmid C
17	101	100.0	6900	4 AAD06045	Aad06045 Plasmid C
18	101	100.0	6956	4 AAD06052	Aad06052 Plasmid C
19	101	100.0	7038	4 AAD06046	Aad06046 Plasmid C
20	101	100.0	7586	6 ABA92644	AbA92644 Cholera t

21	101	100.0	7960	3 AAA59072	Aaa59072 Nucleotid
22	101	100.0	7960	6 ABA94274	AbA94274 Nucleotid
23	101	100.0	7960	10 ADB75120	AdB75120 Plasmid p
24	101	100.0	7960	10 ADF48754	AdF48754 Fibre exp
25	101	100.0	7989	3 AAA59075	Aaa59075 Nucleotid
26	101	100.0	7989	6 ABA94277	AbA94277 Nucleotid
27	101	100.0	7989	10 ADB75123	AdB75123 Plasmid p
28	101	100.0	7989	10 ADF48757	AdF48757 Fibre exp
29	101	100.0	9626	12 ADM97793	AdM97793 Plasmid p
30	100	99.0	1353	10 ADC26320	AdC26320 Human NOV
31	100	99.0	1353	12 ADM35637	AdM35637 Novel hum
32	100	99.0	1353	12 ADO42484	AdO42484 Human NOV
33	100	99.0	1420	8 ACC62251	Acc62251 Human NOV
34	100	99.0	1461	8 ACC62237	Acc62237 Human NOV
35	100	99.0	1770	10 ADJ94793	AdJ94793 Novel NOV
36	100	99.0	1772	10 ADJ94791	AdJ94791 Novel NOV
37	100	99.0	1772	10 ADJ94795	AdJ94795 Novel NOV
38	100	99.0	1822	10 ACD19336	AcD19336 cDNA enco
39	100	99.0	1822	10 ACD19334	AcD19334 cDNA enco
40	100	99.0	6498	10 AAD62465	Aad62465 Human MCH
41	100	99.0	6498	10 ADH53343	AdH53343 pCDNA3 P1
42	100	99.0	7231	3 AAA59090	Aaa59090 Nucleotid
43	100	99.0	7231	6 ABA94286	AbA94286 Nucleotid
44	100	99.0	7231	10 ADB75132	AdB75132 Plasmid p
45	100	99.0	7231	10 ADF48774	AdF48774 Adenoviru

#### ALIGNMENTS

##### RESULT 1

ADA05887

ID ADA05887 standard; cDNA; 2017 BP.

XX AC ADA05887;

XX AC ADA05887;

DT 06-NOV-2003 (first entry)

XX DE

Human NOV55b encoding cDNA SEQ ID NO:247.

human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
immunomodulator; cytostatic; neurotropic; neuroprotective;  
antiparkinsonian; antilipaemic; gene therapy; human disease;  
metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 134..1840

XX FT /\*tag= a

XX FT /product= "NOV55b"

XX WC2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 05-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 18-OCT-2001; 2001US-0330309P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0339266P.



PA (LILL//) LI L.  
PA (GUOX//) GUO X.  
PA (PATT//) PATTURAJAN M.  
PA (SPYT//) SPYTEK K A.  
PA (EDIN//) EDINGER S R.  
PA (ELLE//) ELLERMAN K.  
PA (MALY//) MALYANKAR U M.  
PA (ORTT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B D.  
PA (ANDE//) ANDERSON D W.  
PA (ZHON//) ZHONG M.  
PA (CATT//) CATTERTON E.  
PA (JIWW//) JI W.  
PA (MILL//) MILLER C E.  
PA (RST//) RASTELLI L.  
PA (STON//) STONE D J.  
PA (PENA//) PENNA C E A.  
PA (SHEN//) SHENOY S G.  
PA (SHIM//) SHIMKETS R A.  
PA (ROTH//) ROTHENBERG M E.  
PA (LEAC//) LEACH M D.  
PA (AGEE//) AGEE M L.  
PA (BERG//) BERGHS C.  
PA (DIP//) DIPIPPO V A.  
PA (EISE//) EISEN A.  
PA (GANG//) GANGOLLI E A.  
PA (RIEG//) RIEGER D K.  
PA (SPAD//) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR P-PSDB; ADN63051.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 20; SEQ ID NO 247; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents DNA encoding a human NOVX protein.

XX SQ Sequence 2017 BP; 351 A; 625 C; 599 G; 441 T; 0 U; 1 Other;  
Query Match 100.0%; Score 101; DB 12; Length 2017;  
Best Local Similarity 100.0%; Pred. No. 2.4e-24;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCGCGCGTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60  
Db 1839 AAGCGCGCGTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 1898  
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 101  
Db 1899 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 1939  
RESULT 3  
ID ADO42504 standard; cDNA; 2017 BP.  
XX  
AC ADO42504;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Human NOVX polynucleotide #177.  
XX  
KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;  
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;  
KW scleroderma; hypertension; haemophilia;  
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;  
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;  
KW cancer-associated cachexia; multiple sclerosis; fertility.  
OS Homo sapiens.  
XX  
FN US2004058338-A1.  
XX  
PD 25-MAR-2004.  
XX  
PF 02-DEC-2002; 2002US-00307817.  
XX  
PR 03-DEC-2001; 2001US-0336881P.  
PR 05-DEC-2001; 2001US-0336820P.  
PR 07-DEC-2001; 2001US-0338285P.  
PR 07-DEC-2001; 2001US-0338318P.  
PR 10-DEC-2001; 2001US-0338989P.  
PR 10-DEC-2001; 2001US-0339022P.  
PR 11-DEC-2001; 2001US-0339314P.  
PR 11-DEC-2001; 2001US-0339516P.  
PR 11-DEC-2001; 2001US-0339517P.  
PR 11-DEC-2001; 2001US-0339611P.  
PR 12-DEC-2001; 2001US-0340981P.  
PR 12-DEC-2001; 2001US-0341346P.  
PR 14-DEC-2001; 2001US-0340390P.  
PR 14-DEC-2001; 2001US-0340440P.  
PR 14-DEC-2001; 2001US-0340565P.  
PR 14-DEC-2001; 2001US-0340608P.  
PR 14-DEC-2001; 2001US-0341144P.  
PR 17-DEC-2001; 2001US-0341477P.  
PR 17-DEC-2001; 2001US-0341540P.  
PR 18-DEC-2001; 2001US-0341768P.  
PR 20-DEC-2001; 2001US-0342592P.  
PR 20-DEC-2001; 2001US-0344903P.  
PR 01-FEB-2002; 2002US-0353286P.  
PR 01-FEB-2002; 2002US-0353288P.  
PR 26-FEB-2002; 2002US-0359599P.  
PR 26-FEB-2002; 2002US-0359626P.  
PR 26-FEB-2002; 2002US-0359671P.  
PR 27-FEB-2002; 2002US-0359914P.  
PR 27-FEB-2002; 2002US-0359956P.  
PR 28-FEB-2002; 2002US-0360924P.  
PR 28-FEB-2002; 2002US-0360964P.  
PR 28-FEB-2002; 2002US-0361028P.

PR 28-FEB-2002; 2002US-0361256P.  
PR 28-FEB-2002; 2002US-0361264P.  
PR 05-MAR-2002; 2002US-0361770P.  
PR 05-MAR-2002; 2002US-0362230P.  
PR 13-MAR-2002; 2002US-0364181P.  
PR 13-MAR-2002; 2002US-0364238P.  
PR 15-MAR-2002; 2002US-0364978P.  
PR 15-MAR-2002; 2002US-0365025P.  
PR 17-APR-2002; 2002US-0373288P.  
PR 15-MAY-2002; 2002US-0380981P.  
PR 16-MAY-2002; 2002US-0381004P.  
PR 17-MAY-2002; 2002US-0381495P.  
PR 28-MAY-2002; 2002US-0383534P.  
PR 28-MAY-2002; 2002US-0383744P.  
PR 29-MAY-2002; 2002US-0383829P.  
PR 29-MAY-2002; 2002US-0384024P.  
PR 02-JUL-2002; 2002US-0393332P.  
PR 06-AUG-2002; 2002US-0401315P.  
PR 07-AUG-2002; 2002US-0401788P.  
PR 20-AUG-2002; 2002US-0404676P.  
PR 23-AUG-2002; 2002US-0405400P.  
PR 23-AUG-2002; 2002US-0405684P.  
PR 23-AUG-2002; 2002US-0405687P.  
PR 23-AUG-2002; 2002US-0405698P.  
PR 26-AUG-2002; 2002US-0406353P.  
XX  
PA (AGEE//) AGEE M L.  
PA (ALSO//) ALSOBROOK J P.  
PA (ANDE//) ANDERSON D W.  
PA (BERG//) BERGHS C.  
PA (BOLD//) BOLDOG F L.  
PA (BURG//) BURGESS C E.  
PA (CAIT//) CAITERTON E.  
PA (DIP//) DIPIPPO V A.  
PA (EDIN//) EDINGER S R.  
PA (EISE//) EISEN A.  
PA (ELLE//) ELLERMAN K.  
PA (GANG//) GANGOLLI E A.  
PA (GERL//) GERLACH V.  
PA (GORM//) GORMAN L.  
PA (ROTH//) ROTHBERG B G.  
PA (GUOX//) GUO X S.  
PA (HERR//) HERRMANN J L.  
PA (HALV//) HALVORSEN Y.  
PA (JIW//) JI W.  
PA (KEKU//) KEKUDA R.  
PA (KHRA//) KHRAMTSOV N V.  
PA (LARO//) LAROCHELLE W J.  
PA (LEPL//) LEPLEY D M.  
PA (LILL//) LI L.  
PA (MACD//) MACDOUGALL J R.  
PA (MILL//) MILLER C E.  
PA (ORTT//) ORT T.  
PA (PADI//) PADIGARU M.  
PA (PENA//) PENA C E A.  
PA (PEYM//) PEYMAN J A.  
PA (RIEG//) RIEGER D K.  
PA (ROTH//) ROTHENBERG M E.  
PA (SHEN//) SHENOY S G.  
PA (SMIT//) SMITHSON G.  
PA (SPAD//) SPADERNA S K.  
PA (SPYT//) SPYTEK K A.  
PA (STON//) STONE D J.  
PA (TAUP//) TAUPIER R J.  
PA (VERN//) VERNET C A M.  
PA (VOSS//) VOSS E Z.  
PA (ZHON//) ZHONG M.  
XX  
PI Agee ML, Alsbrook JP, Anderson DW, Berghs C, Boldog FL;  
PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;  
PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG,  
PI Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;  
PI Larochele WJ, Lepley DM, Li L, MacDougall JR, Miller CE, Ort T;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Rieger DK;  
PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;  
PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;  
XX WPI; 2004-268786/25.  
DR P-PSDB; ADO42505.  
XX  
PT New human NOVX polypeptides and nucleic acid molecules, useful for  
PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,  
PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or  
PT scleroderma.  
XX  
PS Claim 20; SEQ ID NO 353; 610pp; English.  
XX  
CC The invention relates to human NOVX polypeptides and the polynucleotides  
CC encoding them. The invention also relates to antibodies specific to the  
CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are  
CC useful for manufacturing a medicament for treating a syndrome associated  
CC with a human disease, such as a pathology associated with the NOVX  
CC polypeptide. The sequences are useful for diagnosing, treating or  
CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,  
CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host  
CC disease, scleroderma, hypertension, haemophilia, idiopathic  
CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,  
CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated  
CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used  
CC as hybridisation probes, in chromosome mapping, in tissue typing, in  
CC preventive medicine or in pharmacogenomics. This sequence represents a  
CC human NOVX polynucleotide of the invention.  
XX  
SQ Sequence 2017 BP; 345 A; 613 C; 617 G; 442 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 12; Length 2017;  
Best Local Similarity 100.0%; Pred. No. 2.4e-24;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCGCGCGCTCGAGTCTAGAGGGCGCGGTTAAACCGCTGATCAGCTCGACTGTGCCT 60  
Db 1839 AAGCGCGCGCTCGAGTCTAGAGGGCGCGGTTAAACCGCTGATCAGCTCGACTGTGCCT 1898  
Qy 61 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGTGCC 101  
Db 1899 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGTGCC 1939  
RESULT 4  
ID ADO42500 standard; cDNA; 2017 BP.  
XX ADO42500;  
AC ADO42500;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Human NOVX polynucleotide #175.  
XX  
KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;  
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;  
KW scleroderma; hypertension; haemophilia;  
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;  
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;  
KW cancer-associated cachexia; multiple sclerosis; fertility.  
XX  
OS Homo sapiens.  
XX  
PN US2004058338-A1.  
XX  
PD 25-MAR-2004.  
XX  
XX 02-DEC-2002; 2002US-00307817.  
XX  
XX 03-DEC-2001; 2001US-0336881P.  
PR  
XX 05-DEC-2001; 2001US-0336820P.

```
PR 07-DEC-2001; 2001US-0338285P.
PR 07-DEC-2001; 2001US-0338318P.
PR 10-DEC-2001; 2001US-0338989P.
PR 10-DEC-2001; 2001US-0339022P.
PR 11-DEC-2001; 2001US-0339314P.
PR 11-DEC-2001; 2001US-0339516P.
PR 11-DEC-2001; 2001US-0339517P.
PR 11-DEC-2001; 2001US-0339611P.
PR 12-DEC-2001; 2001US-0340981P.
PR 12-DEC-2001; 2001US-0341346P.
PR 14-DEC-2001; 2001US-0340390P.
PR 14-DEC-2001; 2001US-0340440P.
PR 14-DEC-2001; 2001US-0340565P.
PR 14-DEC-2001; 2001US-0340608P.
PR 14-DEC-2001; 2001US-0341144P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 18-DEC-2001; 2001US-0341768P.
PR 20-DEC-2001; 2001US-0342592P.
PR 31-DEC-2001; 2001US-0344903P.
PR 01-FEB-2002; 2002US-0353286P.
PR 01-FEB-2002; 2002US-0353288P.
PR 26-FEB-2002; 2002US-0359599P.
PR 26-FEB-2002; 2002US-0359626P.
PR 26-FEB-2002; 2002US-0359671P.
PR 27-FEB-2002; 2002US-0359914P.
PR 27-FEB-2002; 2002US-0359956P.
PR 28-FEB-2002; 2002US-0360924P.
PR 28-FEB-2002; 2002US-0360964P.
PR 28-FEB-2002; 2002US-0361028P.
PR 28-FEB-2002; 2002US-0361256P.
PR 28-FEB-2002; 2002US-0361264P.
PR 05-MAR-2002; 2002US-0361770P.
PR 05-MAR-2002; 2002US-0362230P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364238P.
PR 15-MAR-2002; 2002US-0364978P.
PR 15-MAR-2002; 2002US-0365025P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 16-MAY-2002; 2002US-0381004P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 02-JUL-2002; 2002US-0393332P.
PR 06-AUG-2002; 2002US-0401315P.
PR 07-AUG-2002; 2002US-0401788P.
PR 20-AUG-2002; 2002US-0404676P.
PR 23-AUG-2002; 2002US-0405400P.
PR 23-AUG-2002; 2002US-0405684P.
PR 23-AUG-2002; 2002US-0405687P.
PR 23-AUG-2002; 2002US-0405698P.
PR 26-AUG-2002; 2002US-0406353P.
XX (AGEE/) AGEE M L.
PA (ALSO/) ALSOBROOK J P.
PA (ANDE/) ANDERSON D W.
PA (BERG/) BERGHS C.
PA (BOLD/) BOLDOGF L.
PA (BURG/) BURGESS C E.
PA (CATI/) CATTERTON E.
PA (DIP/) DIPIPPO V A.
PA (EDIN/) EDINGER S R.
PA (EISE/) EISEN A.
PA (ELLE/) ELLERMAN K.
PA (GANG/) GANCOLLI E A.
PA (GERL/) GERLACH V.
PA (GORM/) GORMAN L.
PA (ROTH/) ROTHBERG B G.
PA (GUOX/) GUO X S.
PA (HERR/) HERRMANN J L.
PA (HALV/) HALVORSEN Y.
PA (JTWV/) JI W.
PA (KEKU/) KEKUDA R.
PA (KIRA/) KHRAMTSOV N V.
PA (LABO/) LAROUCHELLE W J.
PA (LEPL/) LEFLEY D M.
PA (LILL/) LI L.
PA (MACD/) MACDOUGALL J R.
PA (MILL/) MILLER C E.
PA (ORTT/) ORT T.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENNA C E A.
PA (PEYM/) PEYMAN J A.
PA (RIEG/) RIEGER D K.
PA (ROTH/) ROTHENBERG M E.
PA (SHEN/) SHENOY S G.
PA (SMIT/) SMITHSON G.
PA (SPAD/) SPADERNA S K.
PA (SPYT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (TAUP/) TAUPIER R J.
PA (VERN/) VERNET C A M.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.
XX
PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;
PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;
PI Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;
PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;
PI Padigaru M, Patturajan M, Pena CE, Peyman JA, Rieger DK;
PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;
PI Stone DJ, Taupier RJ, Vernet CM, Voss EZ, Zhong M;
XX
XX WPI; 2004-268786/25.
DR P-PSDB; ADO42501.
XX
XX New human NOVX polypeptides and nucleic acid molecules, useful for
PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
PT scleroderma.
XX
XX Claim 20; SEQ ID NO 349; 610pp; English.
XX
XX The invention relates to human NOVX polypeptides and the polynucleotides
XX encoding them. The invention also relates to antibodies specific to the
XX NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
XX useful for manufacturing a medicament for treating a syndrome associated
XX with a human disease, such as a pathology associated with the NOVX
XX polypeptide. The sequences are useful for diagnosing, treating or
XX preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
XX diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
XX disease, scleroderma, hypertension, haemophilia, idiopathic
XX thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
XX obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
XX cachexia, multiple sclerosis or fertility. The nucleic acids may be used
XX as hybridisation probes, in chromosome mapping, in tissue typing, in
XX preventive medicine or in pharmacogenomics. This sequence represents a
XX human NOVX polynucleotide of the invention.
XX
SQ Sequence 2017 BP; 345 A; 613 C; 617 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 2017;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGCTGCT 60
Db 1839 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGCTGCT 1898
Qy 61 TCTAGTTCAGCCATCTGTTGTTGCTCCCTCCCGTCC 101
|||||
```

Db 1899 TCTAGTTCAGCCATCTGTGTGTTGGCCCTCCCGGTC 1939

RESULT 5

ADO42506

ID ADO42506 standard; cDNA; 2022 BP.

XX AC ADO42506;

XX DT 15-JUL-2004 (first entry)

XX DE Human NOVX polynucleotide #178.

XX KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;

KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;

KW scleroderma; hypertension; haemophilia;

KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;

KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;

XX KW cancer-associated cachexia; multiple sclerosis; fertility.

OS Homo sapiens.

XX PN US2004058338-A1.

XX PD 25-MAR-2004.

XX PF 02-DEC-2002; 2002US-00307817.

XX PR 03-DEC-2001; 2001US-0336881P.

PR 05-DEC-2001; 2001US-0336820P.

PR 07-DEC-2001; 2001US-0338285P.

PR 07-DEC-2001; 2001US-0338318P.

PR 10-DEC-2001; 2001US-0338989P.

PR 10-DEC-2001; 2001US-0339022P.

PR 11-DEC-2001; 2001US-0339314P.

PR 11-DEC-2001; 2001US-0339516P.

PR 11-DEC-2001; 2001US-0339517P.

PR 11-DEC-2001; 2001US-0339611P.

PR 12-DEC-2001; 2001US-0340981P.

PR 12-DEC-2001; 2001US-0341346P.

PR 14-DEC-2001; 2001US-0340390P.

PR 14-DEC-2001; 2001US-0340440P.

PR 14-DEC-2001; 2001US-0340565P.

PR 14-DEC-2001; 2001US-0340608P.

PR 14-DEC-2001; 2001US-0341144P.

PR 17-DEC-2001; 2001US-0341477P.

PR 17-DEC-2001; 2001US-0341540P.

PR 20-DEC-2001; 2001US-0341768P.

PR 20-DEC-2001; 2001US-0342592P.

PR 31-DEC-2001; 2001US-0344903P.

PR 01-FEB-2002; 2002US-0353286P.

PR 01-FEB-2002; 2002US-0353288P.

PR 26-FEB-2002; 2002US-0359599P.

PR 26-FEB-2002; 2002US-0359626P.

PR 26-FEB-2002; 2002US-0359671P.

PR 27-FEB-2002; 2002US-0359914P.

PR 27-FEB-2002; 2002US-0359956P.

PR 28-FEB-2002; 2002US-0360924P.

PR 28-FEB-2002; 2002US-0360964P.

PR 28-FEB-2002; 2002US-0361028P.

PR 28-FEB-2002; 2002US-0361256P.

PR 28-FEB-2002; 2002US-0361264P.

PR 05-MAR-2002; 2002US-0361770P.

PR 05-MAR-2002; 2002US-0362230P.

PR 13-MAR-2002; 2002US-0364181P.

PR 13-MAR-2002; 2002US-0364238P.

PR 15-MAR-2002; 2002US-0364978P.

PR 15-MAR-2002; 2002US-0365025P.

PR 17-APR-2002; 2002US-0373288P.

PR 15-MAY-2002; 2002US-0380981P.

PR 16-MAY-2002; 2002US-0381004P.

PR 17-MAY-2002; 2002US-0381495P.

PR 28-MAY-2002; 2002US-0383534P.

PR 28-MAY-2002; 2002US-0383744P.

PR 29-MAY-2002; 2002US-0383829P.

PR 29-MAY-2002; 2002US-0384024P.

PR 02-JUL-2002; 2002US-0393332P.

PR 06-AUG-2002; 2002US-0401315P.

PR 07-AUG-2002; 2002US-0401788P.

PR 20-AUG-2002; 2002US-0404676P.

PR 23-AUG-2002; 2002US-0405400P.

PR 23-AUG-2002; 2002US-0405684P.

PR 23-AUG-2002; 2002US-0405687P.

PR 23-AUG-2002; 2002US-0405698P.

XX 26-AUG-2002; 2002US-0406353P.

PA (AGEE/) AGEE M L.

PA (ALSO/) ALSOBROOK J P.

PA (ANDE/) ANDERSON D W.

PA (BERG/) BERGHS C.

PA (BOLD/) BOLDOG F L.

PA (BURG/) BURGESS C B.

PA (CATT/) CATTERTON E.

PA (DIPI/) DIPIPPO V A.

PA (EDIN/) EDINGER S R.

PA (EISE/) EISEN A.

PA (ELLE/) ELLERMAN K.

PA (GANG/) GANGOLLI E A.

PA (GERL/) GERLACH V.

PA (GORM/) GORMAN L.

PA (ROTH/) ROTHBERG B G.

PA (GUOX/) GUO X S.

PA (HERR/) HERRMANN J L.

PA (HALV/) HALVORSEN Y.

PA (JIWW/) JI W.

PA (KEKU/) KEKUDA R.

PA (KHRA/) KHRAMTSOV N V.

PA (LARO/) LAROCHELLE W J.

PA (LEPL/) LEPLEY D M.

PA (LILL/) LI L.

PA (MACD/) MACDOUGALL J R.

PA (MILL/) MILLER C E.

PA (ORTT/) ORT T.

PA (PADI/) PADIGARU M.

PA (PATT/) PATTURAJAN M.

PA (PENA/) PENNA C E A.

PA (PEYM/) PEYMAN J A.

PA (RIEG/) RIEGER D K.

PA (ROTH/) ROTHENBERG M E.

PA (SHEN/) SHENOY S G.

PA (SMIT/) SMITHSON G.

PA (SPAD/) SPADERNA S K.

PA (SPYT/) SPYTEK K A.

PA (STON/) STONE D J.

PA (TAUP/) TAUPIER R J.

PA (VERN/) VERNET C A M.

PA (VOSS/) VOSS E Z.

PA (ZHON/) ZHONG M.

XX Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;

PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;

PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;

PI Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;

PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;

PI Padigaru M, Patturajan M, Pena CE, Peyman JA, Rieger DK;

PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;

PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;

XX WPI: 2004-268786/25.

DR P-PSDB; ADO42507.

XX New human NOVX polypeptides and nucleic acid molecules, useful for

PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,

PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or

PT scleroderma.

XX

```
PS Claim 20; SEQ ID NO 355; 610pp; English.
XX
XX The invention relates to human NOVX polypeptides and the polynucleotides
XX encoding them. The invention also relates to antibodies specific to the
XX NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
XX useful for manufacturing a medicament for treating a syndrome associated
XX with a human disease, such as a pathology associated with the NOVX
XX polypeptide. The sequences are useful for diagnosing, treating or
XX preventing a NOVX-associated disorder, e.g., cancer, graft-versus-host
XX diabetes, Alzheimer's disease, Parkinson's disease, cancer, atherosclerosis,
XX disease, scleroderma, hypertension, haemophilia, idiopathic
XX thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
XX obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
XX cachexia, multiple sclerosis or fertility. The nucleic acids may be used
XX as hybridisation probes, in chromosome mapping, in tissue typing, in
XX preventive medicine or in pharmacogenomics. This sequence represents a
XX human NOVX polynucleotide of the invention.
SQ Sequence 2022 BP; 347 A; 614 C; 618 G; 443 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 2022;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
DB 1843 AAGCGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 1902
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 101
DB 1903 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 1943
RESULT 6
RAD10237
ID AAD10237 standard; DNA; 6050 BP.
XX
XX AAD10237;
XX
XX 24-SEP-2001 (first entry)
XX
XX Commercial plasmid vector pVAX1/lacZ.
XX
XX Plasmid; cyclic; replicon; exogenous gene; marker gene;
XX transcription termination; immunostimulatory sequence; ISS; antiviral;
XX non-essential nucleotide; molecular biology application; gene therapy;
XX DNA vaccine; cloning; gene expression; in vitro protein production;
XX cytosstatic; pVAX1/lacZ; cytomegalovirus promoter; lacZ gene;
XX kanamycin resistant; ds.
XX
XX Cytomegalovirus.
XX Unidentified.
XX Chimeric.
XX
XX Key Location/Qualifiers
XX misc_feature 1..112
XX /tag= a
XX /note= "Corresponds to the non-essential nucleotide
XX sequence that have been removed in the novel DNA plasmid
XX vector"
XX misc_feature 5092..5143
XX /tag= b
XX /note= "3, non-coding region of lacZ gene from pUC18 that
XX corresponds to the non-essential nucleotide sequence that
XX have been removed in the novel DNA plasmid vector"
XX misc_feature 5144..5258
XX /tag= c
XX /note= "Ampicillin promoter that corresponds to the non-
XX essential nucleotide sequence that have been removed in
XX the novel DNA plasmid vector"
XX
XX WO200151626-A2.
XX
```

---

```
PD 19-JUL-2001.
XX
XX 09-JAN-2001; 2001WO-US001255.
XX
XX 10-JAN-2000; 2000US-00480879.
XX
XX (ELIM-) ELIM BIOPHARMACEUTICALS INC.
XX
XX Lu X, Sun L, Zhang Y;
XX
XX WPI; 2001-451855/48.
XX
XX New plasmid DNA vectors, useful for most molecular biology applications,
XX e.g. gene therapy, DNA vaccines, cloning and expression of genes, and in
XX the in vitro production of polypeptides and/or proteins.
XX
XX Example 2; Page 39-41; 50pp; English.
XX
XX The present invention relates to plasmid DNA vectors comprising
XX essentially of a replicon and at least one other component selected from
XX promoter, intron, exogenous gene, transcription termination sequence,
XX selectable marker gene, detectable marker gene and an immunostimulatory
XX sequence (ISS), where the non-essential nucleotide sequences have been
XX substantially removed from these vectors. The plasmid DNA vectors are
XX useful in most molecular biology applications, e.g. gene therapy, DNA
XX vaccines, cloning and expression of genes, and in the in vitro production
XX of polypeptides and/or proteins. The present sequence is a commercial
XX plasmid DNA vector pVAX1/lacZ which comprises Cytomegalovirus promoter,
XX the lacZ gene, pUC origin of replication and the kanamycin resistance
XX gene
XX
XX Sequence 6050 BP; 1346 A; 1597 C; 1696 G; 1411 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 5; Length 6050;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
DB 3828 AAGCGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 3887
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 101
DB 3888 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 3928
RESULT 7
AAD06054
ID AAD06054 standard; DNA; 6567 BP.
XX
XX AAD06054;
XX
XX 31-JUL-2001 (first entry)
XX
XX Plasmid E2CLEDAS encoding fusion protein comprising E2C ZFP-ER LBD-TA.
XX
XX Plasmid E2CLEDAS; fusion protein; nucleotide-binding domain; NBD;
XX ligand-binding domain; LBD; transcription regulating domain; TRD;
XX zinc finger protein; ZFP; ligand-activated transcriptional regulator;
XX gene regulation; gene therapy; cell proliferative disorder; cancer;
XX psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
XX E2C ZFP; human; oestrogen receptor; ER; VP16; TA; transactivation domain;
XX cyclic; circular; ds.
XX
XX Unidentified.
XX Homo sapiens.
XX Herpes simplex virus.
XX Cytomegalovirus.
XX Enterobacteria phage T7.
XX Rhesus macaque polyoma virus.
XX Chimeric.
XX
XX WO200130843-A1.
XX
```

```
XX 03-MAY-2001.
PD
XX
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
XX
XX 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS ) NOVARTIS AG.
XX (SCRI ) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
XX domains, useful e.g. in gene therapy of cancer, provides ligand-activated
XX control of gene expression.
XX
XX Claim 25; Page 186-187; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
XX domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
XX (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
XX zinc finger protein (ZFP), or a modular part of it, that interacts
XX specifically with a contiguous sequence of at least 3 nucleotides. The
XX fusion protein functions as a ligand-activated transcriptional regulator.
XX The fusion protein and the nucleic acid encoding it, are used to regulate
XX gene expression, particularly in gene therapy for treating malignant cell
XX proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
XX carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
XX pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
XX fusion protein and its DNA are also useful for treating diseases caused
XX by viruses in humans/plants, genetic and/or acquired diseases. The fusion
XX protein can be designed to target any selected gene (endogenous or
XX exogenous), and can be made to have different selectivity or specificity
XX for endogenous or exogenous ligands. The present sequence is E2CLBDAS
XX construct encoding fusion protein comprising E2C zinc finger protein
XX (ZFP) that binds human erbB-2 target sequence e2c, human oestrogen
XX receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
XX simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
XX transcription activator. The E2CLBDAS construct is based on plasmid
XX pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
XX and simian virus 40 (SV40)
XX
XX Sequence 6567 BP; 1523 A; 1742 C; 1683 G; 1619 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 4; Length 6567;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-24;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAGCGCGCGTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 60
XX
XX 2171 AAGCGCGCGTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 2230
XX
XX 61 TCTAGTTGCCAGCCATCTGTGTTTTCGCCCTCCCGCGTGCC 101
XX
XX 2231 TCTAGTTGCCAGCCATCTGTGTTTTCGCCCTCCCGCGTGCC 2271
XX
XX
XX RESULT 8
XX AAD06049
XX ID AAD06049 standard; DNA; 6623 BP.
XX
XX AC AAD06049;
XX
XX 31-JUL-2001 (first entry)
XX
XX Plasmid C7LBDAS encoding fusion protein comprising C7 ZFP-ER LBD-TA.
XX
XX Plasmid C7LBDAS; fusion protein; nucleotide-binding domain; NBD;
XX ligand-binding domain; LBD; transcription regulating domain; TRD;
XX zinc finger protein; ZFP; ligand-activated transcriptional regulator;
XX
```

```
KW gene regulation; gene therapy; cell proliferative disorder; cancer;
KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW transactivation domain; cyclic; circular; ds.
XX
XX Mus sp.
XX
XX Homo sapiens.
XX
XX Herpes simplex virus.
XX
XX Cytomegalovirus.
XX
XX Enterobacteria phage T7.
XX
XX Rhesus macaque polyoma virus.
XX
XX Unidentified.
XX
XX Chimeric.
XX
XX WO200130843-A1.
XX
XX 03-MAY-2001.
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
XX
XX 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS ) NOVARTIS AG.
XX (SCRI ) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
XX domains, useful e.g. in gene therapy of cancer, provides ligand-activated
XX control of gene expression.
XX
XX Claim 25; Page 176-178; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
XX domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
XX (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
XX zinc finger protein (ZFP), or a modular part of it, that interacts
XX specifically with a contiguous sequence of at least 3 nucleotides. The
XX fusion protein functions as a ligand-activated transcriptional regulator.
XX The fusion protein and the nucleic acid encoding it, are used to regulate
XX gene expression, particularly in gene therapy for treating malignant cell
XX proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
XX carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
XX pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
XX fusion protein and its DNA are also useful for treating diseases caused
XX by viruses in humans/plants, genetic and/or acquired diseases. The fusion
XX protein can be designed to target any selected gene (endogenous or
XX exogenous), and can be made to have different selectivity or specificity
XX for endogenous or exogenous ligands. The present sequence is C7LBDAS
XX construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
XX which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
XX receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
XX simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
XX transcription activator. The C7LBDAS construct is based on plasmid
XX pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
XX and simian virus 40 (SV40)
XX
XX Sequence 6623 BP; 1530 A; 1754 C; 1703 G; 1636 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 4; Length 6623;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-24;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAGCGCGCGTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 60
XX
XX 2227 AAGCGCGCGTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 2286
XX
XX 61 TCTAGTTGCCAGCCATCTGTGTTTTCGCCCTCCCGCGTGCC 101
XX
XX 2287 TCTAGTTGCCAGCCATCTGTGTTTTCGCCCTCCCGCGTGCC 2327
XX
XX
```



RESULT 9  
 AAD06055  
 ID AAD06055 standard; DNA; 6639 BP.  
 AC  
 AC AAD06055;  
 DT 31-JUL-2001 (first entry)  
 XX  
 XX Plasmid E2CLBDBS encoding fusion protein comprising E2C ZFP-ER LBD-TA.  
 DE  
 DE Plasmid E2CLBDBS; fusion protein; nucleotide-binding domain; NBD;  
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;  
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;  
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;  
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;  
 KW E2C ZFP; human; oestrogen receptor; ER; VP16; TA; transactivation domain;  
 KW cyclic; circular; ds.  
 XX  
 OS Unidentified.  
 OS Homo sapiens.  
 OS Herpes simplex virus.  
 OS Cytomegalovirus.  
 OS Enterobacteria phage T7.  
 OS Rhesus macaque polyoma virus.  
 OS Chimeric.  
 XX  
 XX WO200130843-A1.  
 XX  
 XX PD 03-MAY-2001.  
 XX  
 XX PF 23-OCT-2000; 2000WO-EP010430.  
 XX  
 XX PR 25-OCT-1999; 99US-00433042.  
 XX  
 XX PR 02-JUN-2000; 2000US-00586625.  
 XX  
 XX PA (NOVS ) NOVARTIS AG.  
 XX (SCRI ) SCRIPPS RES INST.  
 XX  
 XX PI Barbas CF, Kadan M, Beerli R;  
 XX WPI; 2001-308618/32.  
 XX  
 XX DR  
 XX New fusion protein containing nucleotide-binding and ligand-binding  
 PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated  
 PT control of gene expression.  
 XX  
 XX PS Claim 25; Page 188-189; 218pp; English.  
 XX  
 CC The invention relates to fusion protein comprising a nucleotide-binding  
 CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor  
 CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl  
 CC zinc finger protein (ZFP), or a modular part of it, that interacts  
 CC specifically with a contiguous sequence of at least 3 nucleotides. The  
 CC fusion protein functions as a ligand-activated transcriptional regulator.  
 CC The fusion protein and the nucleic acid encoding it, are used to regulate  
 CC gene expression, particularly in gene therapy for treating malignant cell  
 CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
 CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
 CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
 CC fusion protein and its DNA are also useful for treating diseases caused  
 CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
 CC protein can be designed to target any selected gene (endogenous or  
 CC exogenous), and can be made to have different selectivity or specificity  
 CC for endogenous or exogenous ligands. The present sequence is E2CLBDBS  
 CC construct encoding fusion protein comprising E2C zinc finger protein  
 CC (ZFP) that binds human erbB-2 target sequence e2c, human oestrogen  
 CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes  
 CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as  
 CC transcription activator. The E2CLBDBS construct is based on plasmid  
 CC pcDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7  
 CC and simian virus 40 (SV40)

XX SQ Sequence 6639 BP; 1546 A; 1749 C; 1718 G; 1626 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 4; Length 6639;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-24;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCT 60  
 DB 2243 AAGCGCGCGCTCGAGTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCT 2302  
 QY 61 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGCC 101  
 DB 2303 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGCC 2343  
 RESULT 10  
 AAD06051  
 ID AAD06051 standard; DNA; 6695 BP.  
 XX  
 AC AAD06051;  
 XX  
 XX 31-JUL-2001 (first entry)  
 XX  
 DE Plasmid C7LBDBS encoding fusion protein comprising C7 ZFP-ER LBD-TA.  
 XX  
 XX Plasmid C7LBDBS; fusion protein; nucleotide-binding domain; NBD;  
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;  
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;  
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;  
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;  
 KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;  
 KW transactivation domain; cyclic; circular; ds.  
 XX  
 OS Mus sp.  
 OS Homo sapiens.  
 OS Herpes simplex virus.  
 OS Cytomegalovirus.  
 OS Enterobacteria phage T7.  
 OS Rhesus macaque polyoma virus.  
 OS Unidentified.  
 OS Chimeric.  
 XX  
 XX WO200130843-A1.  
 XX  
 XX PD 03-MAY-2001.  
 XX  
 XX PF 23-OCT-2000; 2000WO-EP010430.  
 XX  
 XX PR 25-OCT-1999; 99US-00433042.  
 XX  
 XX PR 02-JUN-2000; 2000US-00586625.  
 XX  
 XX PA (NOVS ) NOVARTIS AG.  
 XX (SCRI ) SCRIPPS RES INST.  
 XX  
 XX PI Barbas CF, Kadan M, Beerli R;  
 XX WPI; 2001-308618/32.  
 XX  
 XX DR  
 XX New fusion protein containing nucleotide-binding and ligand-binding  
 PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated  
 PT control of gene expression.  
 XX  
 XX PS Claim 25; Page 180-182; 218pp; English.  
 XX  
 CC The invention relates to fusion protein comprising a nucleotide-binding  
 CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor  
 CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl  
 CC zinc finger protein (ZFP), or a modular part of it, that interacts  
 CC specifically with a contiguous sequence of at least 3 nucleotides. The  
 CC fusion protein functions as a ligand-activated transcriptional regulator.  
 CC The fusion protein and the nucleic acid encoding it, are used to regulate  
 CC gene expression, particularly in gene therapy for treating malignant cell  
 CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
 CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
 CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
 CC fusion protein and its DNA are also useful for treating diseases caused  
 CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
 CC protein can be designed to target any selected gene (endogenous or  
 CC exogenous), and can be made to have different selectivity or specificity  
 CC for endogenous or exogenous ligands. The present sequence is E2CLBDBS  
 CC construct encoding fusion protein comprising E2C zinc finger protein  
 CC (ZFP) that binds human erbB-2 target sequence e2c, human oestrogen  
 CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes  
 CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as  
 CC transcription activator. The E2CLBDBS construct is based on plasmid  
 CC pcDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7  
 CC and simian virus 40 (SV40)

CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
 CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
 CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
 CC fusion protein and its DNA are also useful for treating diseases caused  
 CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
 CC protein can be designed to target any selected gene (endogenous or  
 CC exogenous), and can be made to have different selectivity or specificity  
 CC for endogenous or exogenous ligands. The present sequence is C7LBDBS  
 CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)  
 CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen  
 CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes  
 CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as  
 CC transcription activator. The C7LBDBS construct is based on plasmid  
 CC pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7  
 CC and simian virus 40 (SV40)

XX SQ Sequence 6695 BP; 1553 A; 1762 C; 1737 G; 1643 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 6695;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-24;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60  
 Db 2299 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 2358

Qy 61 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 101  
 Db 2359 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 2399

RESULT 11

AAAD06057  
 ID AAD06057 standard; DNA; 6695 BP.

XX AC AAD06057;  
 XX 31-JUL-2001 (first entry)

XX Plasmid C7LBDBSG400V encoding fusion protein.

XX Plasmid C7LBDBSG400V; fusion protein; nucleotide-binding domain; NBD;  
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;  
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;  
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;  
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;  
 KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; mutant; mutein;  
 KW VP16; TA; transactivation domain; cyclic; circular; ds.

XX Mus sp.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Herpes simplex virus.  
 OS Cytomegalovirus.  
 OS Enterobacteria phage T7.  
 OS Rhesus macaque polyoma virus.  
 OS Unidentified.

XX WO200130843-A1.  
 XX 03-MAY-2001.  
 XX 23-OCT-2000; 2000WO-EP010430.  
 XX 25-OCT-1999; 95US-00433042.  
 XX 02-JUN-2000; 2000US-00586625.  
 XX (NOVS ) NOVARTIS AG.  
 XX (SCRI ) SCRIPPS RES INST.  
 XX Barbas CF, Kadan M, Beerli R;

DR WPI; 2001-308618/32.

XX New fusion protein containing nucleotide-binding and ligand-binding  
 PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated  
 PT control of gene expression.

XX Claim 25; Page 191-193; 218pp; English.

XX The invention relates to fusion protein comprising a nucleotide-binding  
 CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor  
 CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl  
 CC zinc finger protein (ZFP), or a modular part of it, that interacts  
 CC specifically with a contiguous sequence of at least 3 nucleotides. The  
 CC fusion protein functions as a ligand-activated transcriptional regulator.  
 CC The fusion protein and the nucleic acid encoding it, are used to regulate  
 CC gene expression, particularly in gene therapy for treating malignant cell  
 CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
 CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
 CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
 CC fusion protein and its DNA are also useful for treating diseases caused  
 CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
 CC protein can be designed to target any selected gene (endogenous or  
 CC exogenous), and can be made to have different selectivity or specificity  
 CC for endogenous or exogenous ligands. The present sequence is C7LBDBSG400V  
 CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)  
 CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen  
 CC receptor LBD fragment containing G400V mutation and Herpes simplex virus  
 CC VP16 transactivation domain (TA). The ZFP serves as NBD and VP16 TA domain  
 CC functions as transcription activator. The C7LBDBSG400V construct is based  
 CC on plasmid pCDNA3.1 that comprises sequences from cytomegalovirus,  
 CC bacteriophage T7 and simian virus 40 (SV40)

XX SQ Sequence 6695 BP; 1553 A; 1762 C; 1736 G; 1644 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 6695;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-24;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60  
 Db 2299 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 2358

Qy 61 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 101  
 Db 2359 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 2399

RESULT 12

AAAD06058  
 ID AAD06058 standard; DNA; 6695 BP.

XX AC AAD06058;  
 XX 31-JUL-2001 (first entry)

XX Plasmid C7LBDBSG521R encoding fusion protein.

XX Plasmid C7LBDBSG521R; fusion protein; nucleotide-binding domain; NBD;  
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;  
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;  
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;  
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;  
 KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; mutant; mutein;  
 KW VP16; TA; transactivation domain; cyclic; circular; ds.

XX Mus sp.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Herpes simplex virus.  
 OS Cytomegalovirus.  
 OS Enterobacteria phage T7.  
 OS Rhesus macaque polyoma virus.  
 OS Unidentified.

KW	Plasmid C7LBDAL; fusion protein; nucleotide-binding domain; NBD;
KW	ligand-binding domain; LBD; transcription regulating domain; TRD;
KW	zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW	gene regulation; gene therapy; cell proliferative disorder; cancer;
KW	psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW	C7 ZFP; Zif268; murine; human; oestrogen receptor; ER, VP16; TA;
KW	transactivation domain; cyclic; circular; ds.
XX	Mus sp.
OS	Homo sapiens.
OS	Herpes simplex virus.
OS	Cytomegalovirus.
OS	Enterobacteria phage T7.
OS	Rhesus macaque polyoma virus.
OS	Unidentified.
OS	Chimeric.
XX	
PN	WO200130843-A1.
XX	
PD	03-MAY-2001.
PF	23-OCT-2000; 2000MO-EPO10430.
PR	25-OCT-1999; 99US-00433042.
PR	02-JUN-2000; 2000US-00586625.
XX	(NOVS ) NOVARTIS AG.
PA	(SCRI ) SCRIPPS RES INST.
PI	Barbas CF, Kadan M, Beerli R;
XX	WPI; 2001-308618/32.
DR	
XX	New fusion protein containing nucleotide-binding and ligand-binding
PT	domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT	control of gene expression.
XX	
PS	Claim 25; Page 174-176; 218pp; English.
XX	
CC	The invention relates to fusion protein comprising a nucleotide-binding
CC	domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC	(ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC	zinc finger protein (ZFP), or a modular part of it, that interacts
CC	specifically with a contiguous sequence of at least 3 nucleotides. The
CC	fusion protein functions as a ligand-activated transcriptional regulator.
CC	The fusion protein and the nucleic acid encoding it, are used to regulate
CC	gene expression, particularly in gene therapy for treating malignant cell
CC	proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC	carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC	pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC	fusion protein and its DNA are also useful for treating diseases caused
CC	by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC	protein can be designed to target any selected gene (endogenous or
CC	xenogenous), and can be made to have different selectivity or specificity
CC	for endogenous or exogenous ligands. The present sequence is C7LBDAL
CC	construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC	which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC	receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
CC	simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
CC	transcription activator. The C7LBDAL construct is based on plasmid
CC	pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
CC	and simian virus 40 (SV40)
XX	
SQ	Sequence 6746 BP; 1557 A; 1788 C; 1741 G; 1660 T; 0 U; 0 Other;
	Query Match 100.0%; Score 101; DB 4; Length 6746;
	Best Local Similarity 100.0%; Pred. No. 3.3e-24;
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 AACGGCGGTCAGTCTAGAGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
Db	2350 AACGGCGGTCAGTCTAGAGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 240

```
QY      61 TCTAGTGGCCAGCCATCTGTTGTTGCGCCCTCCCGTGC 101
Db      2410 TCTAGTGGCCAGCCATCTGTTGTTGCGCCCTCCCGTGC 2450

RESULT 14
AAD06050
ID      AAD06050 standard; DNA; 6818 BP.
XX
AC      AAD06050;
XX
DT      31-JUL-2001 (first entry)
XX
DE      Plasmid C7LBDL encoding fusion protein comprising C7 ZFP-ER LBD-TA.
XX
KW      Plasmid C7LBDL; fusion protein; nucleotide-binding domain; NBD;
KW      ligand-binding domain; LBD; transcription regulating domain; TRD;
KW      zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW      gene regulation; gene therapy; cell proliferative disorder; cancer;
KW      psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW      C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW      transactivation domain; cyclic; circular; ds.
XX
OS      Mus sp.
OS      Homo sapiens.
OS      Herpes simplex virus.
OS      Cytomegalovirus.
OS      Enterobacteria phage T7.
OS      Rhesus macaque polyoma virus.
OS      Unidentified.
OS      Chimeric.
XX
FN      WO200130843-A1.
XX
PD      03-MAY-2001.
XX
PF      23-OCT-2000; 2000WO-EP010430.
XX
PR      25-OCT-1999; 99US-00433042.
XX      02-JUN-2000; 2000US-00586625.
XX
PA      (NOVS ) NOVARTIS AG.
PA      (SCRI ) SCRIPPS RES INST.
XX
PI      Barbas CF, Kadan M, Beerli R;
XX      WPI; 2001-308618/32.
XX
PT      New fusion protein containing nucleotide-binding and ligand-binding
PT      domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT      control of gene expression.
XX
PS      Claim 25; Page 178-180; 218pp; English.
XX
CC      The invention relates to fusion protein comprising a nucleotide-binding
CC      domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC      (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC      zinc finger protein (ZFP), or a modular part of it, that interacts
CC      specifically with a contiguous sequence of at least 3 nucleotides. The
CC      fusion protein functions as a ligand-activated transcriptional regulator.
CC      The fusion protein and the nucleic acid encoding it, are used to regulate
CC      gene expression, particularly in gene therapy for treating malignant cell
CC      proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC      carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC      pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC      fusion protein and its DNA are also useful for treating diseases caused
CC      by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC      protein can be designed to target any selected gene (endogenous or
CC      exogenous), and can be made to have different selectivity or specificity
CC      for endogenous or exogenous ligands. The present sequence is C7LBDL
CC      construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC      which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC      receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
```

```
CC      simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
CC      transcription activator. The C7LBDL construct is based on plasmid
CC      pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
CC      and simian virus 40 (SV40)
XX
SQ      Sequence 6818 BP; 1580 A; 1796 C; 1775 G; 1667 T; 0 U; 0 Other;
      Query Match      100.0%; Score 101; DB 4; Length 6818;
      Best Local Similarity 100.0%; Pred. No. 3.3e-24;
      Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAGCGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGACTGTGCCT 60
Db      2422 AAGCGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGACTGTGCCT 2461
QY      61 TCTAGTGGCCAGCCATCTGTTGTTGCGCCCTCCCGTGC 101
Db      2482 TCTAGTGGCCAGCCATCTGTTGTTGCGCCCTCCCGTGC 2522

RESULT 15
AAD06044
ID      AAD06044 standard; DNA; 6828 BP.
XX
AC      AAD06044;
XX
DT      31-JUL-2001 (first entry)
XX
DE      Plasmid 2C7LBDAS encoding fusion protein comprising 2C7 ZFP-ER LBD-TA.
XX
KW      Plasmid 2C7LBDAS; fusion protein; nucleotide-binding domain; NBD;
KW      ligand-binding domain; LBD; transcription regulating domain; TRD;
KW      zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW      gene regulation; gene therapy; cell proliferative disorder; cancer;
KW      psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW      2C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW      transactivation domain; cyclic; circular; ds.
XX
OS      Mus sp.
OS      Homo sapiens.
OS      Herpes simplex virus.
OS      Cytomegalovirus.
OS      Enterobacteria phage T7.
OS      Rhesus macaque polyoma virus.
OS      Unidentified.
OS      Chimeric.
XX
FN      WO200130843-A1.
XX
PD      03-MAY-2001.
XX
PF      23-OCT-2000; 2000WO-EP010430.
XX
PR      25-OCT-1999; 99US-00433042.
XX      02-JUN-2000; 2000US-00586625.
XX
PA      (NOVS ) NOVARTIS AG.
PA      (SCRI ) SCRIPPS RES INST.
XX
PI      Barbas CF, Kadan M, Beerli R;
XX      WPI; 2001-308618/32.
XX
PT      New fusion protein containing nucleotide-binding and ligand-binding
PT      domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT      control of gene expression.
XX
PS      Claim 25; Page 168-170; 218pp; English.
XX
CC      The invention relates to fusion protein comprising a nucleotide-binding
CC      domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC      (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC      zinc finger protein (ZFP), or a modular part of it, that interacts
CC      specifically with a contiguous sequence of at least 3 nucleotides. The
CC      fusion protein functions as a ligand-activated transcriptional regulator.
CC      The fusion protein and the nucleic acid encoding it, are used to regulate
CC      gene expression, particularly in gene therapy for treating malignant cell
CC      proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC      carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC      pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC      fusion protein and its DNA are also useful for treating diseases caused
CC      by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC      protein can be designed to target any selected gene (endogenous or
CC      exogenous), and can be made to have different selectivity or specificity
CC      for endogenous or exogenous ligands. The present sequence is C7LBDL
CC      construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC      which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC      receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
```

CC specifically with a contiguous sequence of at least 3 nucleotides. The  
CC fusion protein functions as a ligand-activated transcriptional regulator.  
CC The fusion protein and the nucleic acid encoding it, are used to regulate  
CC gene expression, particularly in gene therapy for treating malignant cell  
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
CC fusion protein and its DNA are also useful for treating diseases caused  
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
CC protein can be designed to target any selected gene (endogenous or  
CC exogenous), and can be made to have different selectivity or specificity  
CC for endogenous or exogenous ligands. The present sequence is 2C7LBDAS  
CC construct encoding fusion protein comprising 2C7 zinc finger protein  
CC (ZFP) which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen  
CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes  
CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as  
CC transcription activator. The 2C7LBDAS construct is based on plasmid  
CC pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7  
CC and simian virus 40 (SV40)

XX  
SQ Sequence 6828 BP; 1595 A; 1816 C; 1746 G; 1681 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 6828;  
Best Local Similarity 100.0%; Pred. No. 3.3e-24;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCCGCTGATCAGCCCTCGACTGTGCCT 60  
Db 2432 AAGCGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCCGCTGATCAGCCCTCGACTGTGCCT 2491  
QY 61 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGCGGCC 101  
Db 2492 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGCGGCC 2532

Search completed: July 14, 2005, 07:01:43  
Job time : 143.448 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_3944\_4044

Perfect score: 101

Sequence: 1 aagcgccgcgcgtagttag.....gttgcccccctcccggtgcc 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl1:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	99	98.0	378	7	CF315931
2	91	90.1	605	7	CK719567
3	84	83.2	400	7	CK850860
C 4	68.2	67.5	132	9	CR074510
C 5	68.2	67.5	286	9	CR083191
6	66	65.3	295	7	CN778129
7	66	65.3	219	5	BM888450
8	66	65.3	521	5	BM887817
C 9	65.6	65.0	233	9	CR154962
C 10	65	64.4	166	9	CR092687
11	64.4	63.8	534	5	BM887701
12	64.2	63.6	329	9	CG632479
13	63.4	62.8	600	5	BM887768
C 14	61.8	61.2	75	9	CR037248
C 15	61.8	61.2	87	9	CR106833
C 16	61.8	61.2	130	9	BM88352
C 17	61	60.4	234	9	CR070494
C 18	60.2	59.6	330	9	CR006502
C 19	60	59.4	605	5	BM888562
C 20	59.8	59.2	108	9	CR173214
C 21	59.4	58.8	304	9	BM997931
C 22	59	58.4	104	9	CR104210
C 23	59	58.4	110	9	BM982981
C 24	58.6	58.0	323	9	CR100521

C 25	58.2	57.6	158	9	CR018574
C 26	58.2	57.6	197	9	CR014355
C 27	58	57.4	471	4	BM819796
C 28	57.2	56.8	664	9	CR160587
C 29	57.2	56.6	284	9	BM984480
C 30	57	56.4	158	9	CR117924
C 31	56.8	56.2	199	9	CR047320
C 32	56.6	56.0	141	9	CR126132
C 33	56.6	56.0	159	9	CR133954
C 34	56.6	56.0	160	9	CR012517
C 35	56.4	55.8	109	9	CR108493
C 36	56	55.4	99	9	CR009197
C 37	56	55.4	347	9	CR045655
C 38	55.8	55.2	77	9	CR171087
C 39	55.8	55.2	89	9	CR081749
C 40	55.8	55.2	96	9	CR180417
C 41	55.4	54.9	104	9	CR163542
C 42	55.4	54.9	113	9	CR100912
C 43	55.4	54.9	132	9	CR081810
C 44	55.2	54.7	277	9	CR144680
C 45	55	54.5	89	9	CR160035

## ALIGNMENTS

RESULT 1  
CF315931  
LOCUS  
DEFINITION HD--05-A13.b1 OeHDAC1-overexpressing transgenic rice plasmid cDNA  
library (HD) Oryza sativa (japonica cultivar-group) cDNA clone  
HD--05-A13, mRNA sequence.  
ACCESSION CF315931  
VERSION CF315931.1 GI:33687692  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 378)  
AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,  
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.  
TITLE Large-scale Sequencing Analysis of Rice ESTs  
JOURNAL Unpublished (2003)  
COMMENT Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.

## FEATURES

source  
1. 378  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="HD--05-A13"  
/tissue\_type="callus"  
/dev\_stage="proliferated callus on 2N6 media for 2 weeks"  
/lab\_host="E.coli DH10B"  
/clone\_lib="OeHDAC1-overexpressing transgenic rice plasmid  
CDNA library (HD)"  
/notes="vector: PCR4-TOPO; Site 1: EcoRI; Callus was  
treated with ABA(20um) for 1hr. Oligo-capped mRNA was  
reverse transcribed and then used for PCR. mRNA was  
derived from rice Histone Deacetylase overexpression  
line."

## ORIGIN

Query Match 98.0%; Score 99; DB 7; Length 378;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;

```

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 79 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 138
    |||

Qy 63 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 101
    |||
Db 139 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 177
    |||

RESULT 2
CK719567
LOCUS 19817 Swollen Stolon Solanum tuberosum cDNA, mRNA sequence. EST 10-FEB-2004
DEFINITION CK719567
ACCESSION CK719567
VERSION CK719567.1 GI:42511281
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 605)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Laque,M.,
DeKoeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and
Regan,S.
TITLE Generation of ESTs from swollen stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Clones can be requested from BioAtlantech via
bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        Location/Qualifiers
            1..605
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Swollen Stolon"
                /note="Vector: pBluescript II SK(+) XR; Site 1: EcoRI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a greenhouse under natural
                conditions. RNA was isolated from swollen stolon tissue,
                3-10mm in length, which was cut from the tip, to the base
                of swelling."
ORIGIN
Query Match 90.1%; Score 91; DB 7; Length 605;
Best Local Similarity 94.9%; Pred. No. 1e-18; Indels 0; Gaps 0;
Matches 94; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 393 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 452
    |||

Qy 63 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 101
    |||
Db 453 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGAC 491
    |||

RESULT 3
CK850860
LOCUS 10869 Stolon Solanum tuberosum cDNA, mRNA sequence. EST 08-MAR-2004
DEFINITION CK850860
ACCESSION CK850860.1 GI:45239470
VERSION

```

```

KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 400)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Laque,M., De
Koeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.
TITLE Generation of ESTs from stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        Location/Qualifiers
            1..400
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Stolon"
                /note="Vector: pBluescript II SK(+) XR; Site 1: EcoRI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a greenhouse under natural
                conditions. RNA was isolated from stolon tissue."
ORIGIN
Query Match 83.2%; Score 84; DB 7; Length 400;
Best Local Similarity 99.0%; Pred. No. 1.8e-16;
Matches 95; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 224 GCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 283
    |||

Qy 63 TAGTTGCCAGCCA-TCTGTTGTTTGGCCCTCCCTCCCGG 97
    |||
Db 284 TAGTTGCCAGCCA-TCTGTTGTTTGGCCCTCCCTCCCGG 319
    |||

RESULT 4
CR074510/c
LOCUS 132 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP255d22, genomic survey sequence.
ACCESSION CR074510
VERSION CR074510.1 GI:49808100
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 132)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
    source
        Location/Qualifiers
            1..132
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone_lib="MHP255d22"
                /clone_lib="MHP"
ORIGIN

```



```

Query Match          67.5%; Score 68.2; DB 9; Length 132;
Best Local Similarity 90.1%; Pred. No. 1.9e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 21 AGGCCCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
Db 93 ACGACCCCATGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 34

Qy 81 TGTTCGCCCTCCCGCGTCC 101
Db 33 TGTTCGCCCTCCCGCGTCC 13

RESULT 5
CR083191/c
LOCUS
DEFINITION
Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP263n05, genomic survey sequence.
ACCESSION
CR083191
VERSION
CR083191.1 GI:49816780
KEYWORDS
GSS; genome survey sequence; MICER.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 286)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
Location/Qualifiers
source
1..286
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP263n05"
/clone_lib="MHP"

Query Match          67.5%; Score 68.2; DB 9; Length 286;
Best Local Similarity 90.1%; Pred. No. 2.2e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 21 AGGCCCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
Db 88 ACGACCCCATGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 29

Qy 81 TGTTCGCCCTCCCGCGTCC 101
Db 28 TGTTCGCCCTCCCGCGTCC 8

RESULT 6
CN778129
LOCUS
DEFINITION
pgn2c.pk001.h10.f Chicken lymphoid cDNA library (pgn2c) Gallus
gallus cDNA clone pgn2c.pk001.h10.f 3'end of pat.pk0008.d12 5',
mRNA sequence.
ACCESSION
CN778129
VERSION
CN778129.1 GI:47548763
KEYWORDS
Gallus gallus (chicken)
SOURCE
Gallus gallus
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 295)
Morgan,R.W. and Burnside,J.
Chicken ESTs from lymphoid tissue- 3' sequence
Unpublished (2004)
Contact: Robin W. Morgan

Query Match          67.5%; Score 68.2; DB 9; Length 286;
Best Local Similarity 90.1%; Pred. No. 2.2e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 21 AGGCCCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
Db 88 ACGACCCCATGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 29

Qy 81 TGTTCGCCCTCCCGCGTCC 101
Db 28 TGTTCGCCCTCCCGCGTCC 8

RESULT 6
CN778129
LOCUS
DEFINITION
pgn2c.pk001.h10.f Chicken lymphoid cDNA library (pgn2c) Gallus
gallus cDNA clone pgn2c.pk001.h10.f 3'end of pat.pk0008.d12 5',
mRNA sequence.
ACCESSION
CN778129
VERSION
CN778129.1 GI:47548763
KEYWORDS
Gallus gallus (chicken)
SOURCE
Gallus gallus
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 295)
Morgan,R.W. and Burnside,J.
Chicken ESTs from lymphoid tissue- 3' sequence
Unpublished (2004)
Contact: Robin W. Morgan

```

```

University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341
Fax: 302-831-2822
Email: morgan@udel.edu, www.chickest.udel.edu.

FEATURES
Location/Qualifiers
source
1..295
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="pgn2c.pk001.h10.f 3'end of pat.pk0008.d12"
/sex="Male and Female"
/tissue_type="thymus, bursa, spleen, PBL, bone marrow"
/lab_host="E.coli EMDH10B"
/clone_lib="Chicken Lymphoid cDNA library (pgn2c)"
/note="Vector: pCMVSPORT 6"

ORIGIN
Query Match          65.3%; Score 66; DB 7; Length 295;
Best Local Similarity 93.2%; Pred. No. 1.1e-10;
Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 28 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 87
Db 1 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 60

Qy 88 CCTCTCCCGCTGCC 101
Db 61 CCTCTCCCGCTGCC 74

RESULT 7
BM888450
LOCUS
DEFINITION
TM108 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
Clone 104447 5', mRNA sequence.
ACCESSION
BM888450
VERSION
BM888450.1 GI:19272194
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
Expression Profile and Genome Location of cDNA Clones from an
Infant Human Trabecular Meshwork Library
Unpublished (2002)
Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698
Fax: 503-494-6875
Email: wirtzm@ohsu.edu
Seq primer: T7 Reverse.
Location/Qualifiers
source
1..519
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="104447"
/tissue_type="eye"
/cell_type="trabecular meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOP10P"
/clone_lib="Human Trabecular Meshwork cDNA library"
/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
cDNA library made from mRNA isolated from trabecular
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pCDNA3 vector

```

and TPO10F", host cells."

ORIGIN

Query Match 65.3%; Score 66; DB 5; Length 519;  
 Best Local Similarity 93.2%; Pred. No. 1.3e-10;  
 Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 28 GTTTAAACCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 87  
 Db 333 GCTAGAGCTCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 392

QY 88 CCTCTCCCGTGCC 101  
 Db 393 CCTCTCCCGTGCC 406

RESULT 8  
 BM887817 521 bp mRNA linear EST 08-MAR-2002  
 LOCUS TM553 Human Trabecular Meshwork cDNA library Homo sapiens cDNA  
 DEFINITION clone 122060 5', mRNA sequence.

ACCESSION BM887817  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 521)  
 Wirtz, M.K., Samples, J.R., Xu, H., Severson, T. and Acott, T.S.  
 Expression Profile and Genome Location of cDNA Clones from an  
 Infant Human Trabecular Meshwork Library  
 Unpublished (2002)  
 Contact: Wirtz MK  
 Glaucoma Genetics Lab  
 Oregon Health Sciences University  
 3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA  
 Tel: 503-494-4698  
 Fax: 503-494-6875  
 Email: wirtzm@ohsu.edu  
 Seq primer: T7 Reverse  
 High quality sequence stop: 350.

FEATURES  
 Location/Qualifiers  
 1..521  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="122060"  
 /tissue\_type="eye"  
 /cell\_type="trabecular meshwork"  
 /dev\_stage="2 week to 2 year old infants"  
 /lab\_host="TPO10F"  
 /clone\_lib="Human Trabecular Meshwork cDNA library"  
 /note="Vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human  
 cDNA library made from mRNA isolated from trabecular  
 meshwork cells established from eyes from 6 individuals,  
 ages 2 weeks to 2 years. Cells were harvested at passages  
 3 through 6. Invitrogen made a unidirectional cDNA library  
 from the mRNA from the frozen cells using a pcDNA3 vector  
 and TPO10F", host cells."

ORIGIN

Query Match 65.3%; Score 66; DB 5; Length 521;  
 Best Local Similarity 93.2%; Pred. No. 1.3e-10;  
 Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 28 GTTTAAACCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 87  
 Db 348 GCTAGAGCTCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 407

QY 88 CCTCTCCCGTGCC 101  
 Db 408 CCTCTCCCGTGCC 421

and TPO10F", host cells."

ORIGIN

Query Match 65.3%; Score 66; DB 5; Length 519;  
 Best Local Similarity 93.2%; Pred. No. 1.3e-10;  
 Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 28 GTTTAAACCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 87  
 Db 333 GCTAGAGCTCGGTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTTGTC 392

QY 88 CCTCTCCCGTGCC 101  
 Db 393 CCTCTCCCGTGCC 406

RESULT 9  
 CR154962/c 233 bp DNA linear GSS 06-JUL-2004  
 LOCUS Forward strand read from insert in 3'HPRT insertion targeting and  
 DEFINITION chromosome engineering clone MHPPI82j09, genomic survey sequence.  
 CR154962  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE GSS; genome survey sequence; MICER.  
 ORGANISM Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 233)  
 Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,  
 Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,  
 Rogers, J. and Bradley, A.  
 Direct Submission  
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES  
 Location/Qualifiers  
 1..233  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPPI82j09"  
 /clone\_lib="MHPp"

ORIGIN

Query Match 65.0%; Score 65.6; DB 9; Length 233;  
 Best Local Similarity 88.8%; Pred. No. 1.5e-10;  
 Matches 71; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 21 AGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80  
 Db 223 ACGACCCGCTGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 164

QY 81 TGTTTGGCCCTCCCGCTGC 100  
 Db 163 TGTTTGGCCCTCCCGCTGC 144

RESULT 10  
 CR092687/c 166 bp DNA linear GSS 05-JUL-2004  
 LOCUS Forward strand read from insert in 3'HPRT insertion targeting and  
 DEFINITION chromosome engineering clone MHP224h16, genomic survey sequence.  
 CR092687  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE GSS; genome survey sequence; MICER.  
 ORGANISM Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 166)  
 Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,  
 Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,  
 Rogers, J. and Bradley, A.  
 Direct Submission  
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES  
 Location/Qualifiers  
 1..166  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHP224h16"  
 /clone\_lib="MHPp"

ORIGIN

Query Match 64.4%; Score 65; DB 9; Length 166;  
 Best Local Similarity 87.7%; Pred. No. 2.2e-10;

```

Matches 71; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 21 AGGGCCCTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
Db 156 ACGACCCCATCATCGCGCTGATCAGCTCGAGTGTGCTTCTAGTTCGCCAGCCATCTGT 97

Qy 81 TGTGTGCCCTCCCGCGGCC 101
Db 96 TGTGTGCCCTCCCGCGGCC 76

RESULT 11
BM887701 534 bp mRNA linear EST 08-MAR-2002
LOCUS TM304 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
DEFINITION clone 107917 5', mRNA sequence.
ACCESSION BM887701
VERSION BM887701.1 GI:19271430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 534)
REFERENCE Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
AUTHORS Expression Profile and Genome Location of cDNA Clones from an
TITLE Infant Human Trabecular Meshwork Library
JOURNAL Unpublished (2002)
COMMENT Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698
Fax: 503-494-6875
Email: wirtzm@ohsu.edu
Seq primer: 17 Reverse.
Location/Qualifiers
FEATURES
source
1..534
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="107917"
/tissue_type="eye"
/cell_type="trabecular meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOP10P"
/clone_lib="Human Trabecular Meshwork cDNA library"
/note="Vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
cDNA library made from mRNA isolated from trabecular
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pcDNA3 vector
and TPO10P," host cells."

ORIGIN
Query Match 63.8%; Score 64.4; DB 5; Length 534;
Best Local Similarity 91.9%; Pred. No. 4.2e-10;
Matches 68; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 28 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGTC 87
Db 445 GCTAGAGTTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGTC 504

Qy 88 CCCTCCCGCGTGCC 101
Db 505 CCCTCCCGCGTGCC 518

RESULT 12
CG632479 329 bp mRNA linear GSS 02-OCT-2003
LOCUS OST350781 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST350781,
DEFINITION

```

```

mRNA sequence.
CG632479 GI:37456328
VERSION CG632479.1
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 329)
REFERENCE zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
AUTHORS Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
Location/Qualifiers
FEATURES
source
1..329
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="OST350781"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN
Query Match 63.6%; Score 64.2; DB 9; Length 329;
Best Local Similarity 95.7%; Pred. No. 4.5e-10;
Matches 66; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 33 AACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGCCCTC 92
Db 185 AGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGCCCTC 244

Qy 93 CCCCCTGCC 101
Db 245 CCCCCTACC 253

RESULT 13
BM887768 600 bp mRNA linear EST 08-MAR-2002
LOCUS TM397 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
DEFINITION clone 119752 5', mRNA sequence.
ACCESSION BM887768
VERSION BM887768.1 GI:19271512
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 600)
REFERENCE Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
AUTHORS Expression Profile and Genome Location of cDNA Clones from an
TITLE Infant Human Trabecular Meshwork Library
JOURNAL Unpublished (2002)
COMMENT Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698

```

Fax: 503-494-6875  
Email: wirtzm@ohsu.edu  
Seq primer: T7 Reverse  
High quality sequence stop: 400.

## FEATURES

source  
1. .600  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="119752"  
/tissue\_type="eye"  
/cell\_type="trabecular meshwork"  
/dev\_stage="2 week to 2 year old infants"  
/lab\_host="TOP10P"  
/clone\_lib="Human Trabecular Meshwork cDNA library"  
/note="Vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human cDNA library made from mRNA isolated from trabecular meshwork cells established from eyes from 6 individuals, ages 2 weeks to 2 years. Cells were harvested at passages 3 through 6. Invitrogen made a unidirectional cDNA library from the mRNA from the frozen cells using a pcDNA3 vector and TP010P, host cells."

## ORIGIN

Query Match 62.8%; Score 63.4; DB 5; Length 600;  
Best Local Similarity 90.5%; Pred. No. 9e-10; Indels 0; Gaps 0;  
Matches 67; Conservative 0; Mismatches 7;

Qy 28 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 87  
Db 516 GCTAGAGTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 575

Qy 88 CCCTCCCGGTGCC 101  
Db 576 CCCTCCCGGTGCC 589

## RESULT 14

CR037248/c  
LOCUS  
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP31109, genomic survey sequence.  
ACCESSION CR037248  
VERSION CR037248.1 GI:49770303  
KEYWORDS GSS; genome survey sequence; MICER.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.  
Direct Submission  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
Location/Qualifiers  
1. .75  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP31109"  
/clone\_lib="MHP3"

## ORIGIN

Query Match 61.2%; Score 61.8; DB 9; Length 75;  
Best Local Similarity 96.9%; Pred. No. 2.1e-09;  
Matches 63; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 37 CGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCC 96  
Db 74 CGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCC 15

Qy 97 GTGCC 101  
Db 14 GTGCC 10

## RESULT 15

CR106833/c  
LOCUS  
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP37h18, genomic survey sequence.  
ACCESSION CR106833  
VERSION CR106833.1 GI:49854244  
KEYWORDS GSS; genome survey sequence; MICER.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.  
Direct Submission  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
Location/Qualifiers  
1. .87  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP37h18"  
/clone\_lib="MHP3"

## FEATURES

source  
1. .87  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP37h18"  
/clone\_lib="MHP3"

## ORIGIN

Query Match 61.2%; Score 61.8; DB 9; Length 87;  
Best Local Similarity 85.2%; Pred. No. 2.1e-09;  
Matches 69; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 21 AGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGT 80  
Db 86 ACAGCCCATGCGATCGCGATGATCAGCCTCAACTCTGCTTCTAGTTGCCAGCCATCTGT 27

Qy 81 TGTTTGCCCTCCCGGTGCC 101  
Db 26 TGTTTGCCCTCCCGGTGCC 6

Search completed: July 14, 2005, 23:23:05  
Job time : 961.667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)

Title: US-09-482-682-47 COPY 7889 7989

Perfect score:

Sequence: 1 agggttattgtctcatgagc.....gaaaagtgccacctgacgtc 101

Scoring table: IDENTITY NUC

Scoring cubic: IDENTIFICATION  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 2000000000

Post-processing: Minimum Match 0%

100% Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: ★

```
1: gb_ba:*
```

2: gb\_htg:\*

3: gb\_in: \*

4: ggb\_om:★

5: ggb\_ov:★

6: ggb\_pat: \*

7: gb\_ph:★

8: ggb\_p1:\*

9: gfb\_pr:\*

10: gb\_ro:\*

```
11: gb_sts:
```

12: gb\_sy: \*

```
13: ggb_un:*
```

14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	101	100.0	142	6	AR356490	Sequence
C 2	101	100.0	142	6	AR538046	Sequence
C 3	101	100.0	228	6	E00019	DNA coding
C 4	101	100.0	240	1	PMWENDO	MI0139 Plasmid pmM
C 5	101	100.0	251	6	E00018	DNA coding
C 6	101	100.0	251	6	I01644	Sequence 1
C 7	101	100.0	344	11	HUMUT5345	L18624 Human chrom
C 8	101	100.0	400	6	BD195256	Nucleotid
C 9	101	100.0	456	6	E00892	Synthetic D
C 10	101	100.0	456	6	E01156	DNA fragmen
C 11	101	100.0	456	6	E01274	DNA encodin
C 12	101	100.0	456	6	E01302	DNA encodin
C 13	101	100.0	466	6	AX260098	Sequence
C 14	101	100.0	573	6	AX260150	Sequence
C 15	101	100.0	693	6	A43586	Sequence 11
C 16	101	100.0	693	6	AR116755	Sequence
C 17	101	100.0	998	1	AY559171	Pseudomon
C 18	101	100.0	1011	1	SMTEMAQGE	X37254 S.marcscen
C 19	101	100.0	1012	2	CEC11F10	Z92776 Caenorhabdi

## ALIGNMENTS

RESULT 1	AR356490/c	AR356490	142 bp	DNA	linear	PAT 17-AUG-2003
LOCUS		Sequence 2608 from patent US 6593114.				
DEFINITION		AR356490				
ACCESSION		AR356490				
VERSION		AR356490.1	GI:33762574			
KEYWORDS		.				
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		1 (bases 1 to 142)				
AUTHORS		Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.				
TITLE		Staphylococcus aureus polynucleotides and sequences				
JOURNAL		Patent: US 6593114-A 2608 15-JUL-2003;				
FEATURES		Location/Qualifiers				
source		1..142				
		/organism="unknown"				
		/mol_type="genomic DNA"				
ORIGIN						
		Query Match	100.0%;	Score 101;	DB 6;	Length 142;
		Best Local Similarity	100.0%;	Pred. No. 8.7e-20;		
		Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTGAAAAATAAACAAATAG 60				
Db	107	AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTGAAAAATAAACAAATAG 48				
Qy	61	GGGTTCCGCGACATATTTCCCGAAAAGTGCACCTGACGTC 101				
Db	47	GGGTTCCGCGACATATTTCCCGAAAAGTGCACCTGACGTC 7				
RESULT 2	AR538046/c	AR538046	142 bp	DNA	linear	PAT 08-OCT-2004
LOCUS		Sequence 2608 from patent US 6737248.				
DEFINITION		AR538046				
ACCESSION		AR538046				
VERSION		AR538046.1	GI:53929263			
KEYWORDS		.				
SOURCE		Unknown.				

ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 142)  
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.  
TITLE Staphylococcus aureus polynucleotides and sequences  
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;  
FEATURES Location/Qualifiers  
source 1..142  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 142;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60  
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 48  
Qy 61 GGGTTCGGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
Db 47 GGGTTCGGCACATTTCCCGGAAAGTGCCACCTGACGTC 7  
RESULT 3  
E00019/c  
LOCUS DNA coding for Escherichia coli penicillinase.  
DEFINITION E00019  
ACCESSION E00019  
VERSION E00019.1 GI:2168327  
KEYWORDS JP 1981154999-A/2.  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE 1 (bases 1 to 228)  
AUTHORS Uotutaa,G. and Karen,T.  
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA  
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;  
COMMENT UNIV HARVARD  
OS Escherichia coli  
PN JP 1981154999-A/2  
PD 30-NOV-1981  
PF 09-APR-1981 JP 1981052488  
PR 11-APR-1980 US 80 139225  
PI UORUTAA GIRUBATO, KAREN TARUMATSUJI  
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC  
strandedness: Double;  
CC topology: Linear;  
CC anti-sense: No;  
CC \*source: clone=pKT218;  
FH Key Location/Qualifiers  
FH CDS 210..>228  
FT /product='E.coli penicillinase'.  
FEATURES  
source 1..228  
Location/Qualifiers  
1..228  
/organism="Escherichia coli"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:562"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 228;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60  
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 116  
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101

Db 115 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 75  
RESULT 4  
PMOENDO/c  
LOCUS DNA 240 bp linear BCT 26-APR-1993  
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.  
ACCESSION M10199  
VERSION M10199.1 GI:150826  
KEYWORDS  
SOURCE Plasmid pMM110  
ORGANISM Plasmid pMM110  
other sequences; plasmids.  
REFERENCE 1 (bases 1 to 240)  
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.  
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA  
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)  
MEDLINE 85153063  
PUBMED 6397324  
COMMENT Original source text: Plasmid pMM110 DNA.  
FEATURES Location/Qualifiers  
source 1..240  
/organism="Plasmid pMM110"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2599"  
/plasmid="Plasmid pMM110"  
ORIGIN Unreported.  
Query Match 100.0%; Score 101; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60  
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 92  
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
Db 91 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 51  
RESULT 5  
E00018/c  
LOCUS DNA coding for Escherichia coli penicillinase.  
DEFINITION E00018  
ACCESSION E00018  
VERSION E00018.1 GI:2168326  
KEYWORDS JP 1981154999-A/1.  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE 1 (bases 1 to 251)  
AUTHORS Uotutaa,G. and Karen,T.  
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA  
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;  
COMMENT UNIV HARVARD  
OS Escherichia coli  
PN JP 1981154999-A/1  
PD 30-NOV-1981  
PF 09-APR-1981 JP 1981052488  
PR 11-APR-1980 US 80 139225  
PI UORUTAA GIRUBATO, KAREN TARUMATSUJI  
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC  
strandedness: Double;  
CC topology: Linear;  
CC anti-sense: No;  
CC fragment\_type: N-Terminal Fragment;  
CC \*source: clone=pKT241;

```

FH Key Location/Qualifiers
FT CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES
    source Location/Qualifiers
        1..251
        /organism='Escherichia coli'
        /mol_type='genomic DNA'
        /db_xref='taxon:562'
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 251;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 116
    |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 75
    |||||||
RESULT 6
I01644/c
LOCUS 251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
    1 (bases 1 to 251)
    AUTHORS Gilbert,W. and Taimadge,K.
    TITLE Mature protein synthesis
    JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
    President and Fellows of Harvard College; Cambridge, MA
FEATURES
    source Location/Qualifiers
        1..251
        /organism='unknown'
        /mol_type='unassigned DNA'
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 251;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 116
    |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 75
    |||||||
RESULT 7
HUMUT5345
LOCUS 344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;
microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 344)
    Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

```

```

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAACAGGAGGCAAAATGC
Primer B: TTCGGGAATGTCCCGGAAC
32P-label: B Primer
PCR Profile:
    Initial Denaturation: 94C 300sec
    PCR Cycles: 30
    Denaturation: 94C 10sec
    Annealing: 60C 10sec
    Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2
FEATURES
    source Location/Qualifiers
        1..344
        /organism='Homo sapiens'
        /mol_type='genomic DNA'
        /db_xref='taxon:9606'
        /map='8'
        /map_pos='36..224'
        /standard_name='STS UT5345'
        primer_bind 36..60
        primer_bind complement(202..224)
ORIGIN
    Query Match 100.0%; Score 101; DB 11; Length 344;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 141 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 200
    |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 201 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 241
    |||||||
RESULT 8
BD195256/c
LOCUS 400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE
    1 (bases 1 to 400)
    Dillion,P.J., Choi,G.H. and Welch,R.A.
    Nucleotide sequence of Escherichia coli pathogenicity islands
    Patent: JP 2002513277-A 43 08-MAY-2002;
    HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT
    OS Unidentified
    PN JP 2002513277-A/43
    PD 08-MAY-2002
    PF 21-NOV-1997 JP 1998523916
    PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
    PATRICK J DILLON,GIL H CHOI,RODNEY A WELCH
    PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
    Strandedness: Double;
    CC Topology: Linear;
    CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.
FEATURES
  source
    1..400
    /organism='unidentified'
    /mol_type='genomic DNA'
    /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 165 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 106
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
Db 105 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clones=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FH of beta-lactamase
FH RBS 200..203
FH CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
FEATURES
  source
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'
```

```

ORIGIN
/db_xref='taxon:32630'
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH Key Location/Qualifiers
FH promoter 125..170
FH /note='beta lactamase promoter' FT RBS
FH CDS 200..204
FH /product='beta urogastrone'
FH sig_peptide 209..277
FH mat_peptide 278..436
FT /product='beta urogastrone'.
FT Location/Qualifiers
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
```



```
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS
DEFINITION
DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION
E01274
VERSION
E01274.1 GI:2169533
KEYWORDS
JP 1987179398-A/1.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaihara,N.
TITLE
PRODUCTION OF BETA-UROGASTRONE
JOURNAL
Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT
EARTH CHEM CORP LTD
OC Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Location/Qualifiers
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta-urogastron'
FT CDS 209..439
FT /product='beta-urogastron'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
Qy 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS
DEFINITION
DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION
E01302
VERSION
E01302.1 GI:2169561
KEYWORDS
JP 1987190083-A/1.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaihara,N.
TITLE
PRODUCTION OF BETA-UROGASTRONE
JOURNAL
Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT
EARTH CHEM CORP LTD
OC Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Location/Qualifiers
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta-urogastron'
FT CDS 209..439
FT /product='beta-urogastron'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
Qy 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS
DEFINITION
Sequence 60 from Patent WO0172774.
ACCESSION
AX260098
VERSION
AX260098.1 GI:16509129
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1
AUTHORS
Deak,P., Glover,D.M. and Midgley,C.
TITLE
Cell cycle progression proteins
JOURNAL
Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
```

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 221
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
    |||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION      AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1
REFERENCE
AUTHORS      Deak, P., Glover, D.M. and Midgley, C.
TITLE      Cell cycle progression proteins
JOURNAL      Patent: WO 0172774-A 112 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..573
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 296
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 295 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 255
    |||||

RESULT 15
A43586
LOCUS
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION      A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM
Cuphea lanceolata
Cuphea lanceolata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Myrtales; Lythraceae; Cuphea.
1 (bases 1 to 693)
Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
Hoerhke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
Schulte, W., Voeltz, M., Walek, J. and Schell, J.
PROMOTERS
TITLE      Patent: WO 9507357-A 11 16-MAR-1995;
JOURNAL      MAX PLANCK GESELLSCHAFT (DE)
COMMENT      Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 651
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 652 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 692
    |||||

Search completed: July 14, 2005, 14:03:32
Job time : 756.618 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_7889\_7989  
Perfect score: 101  
Sequence: 1 aggtttattgtctcatgacg.....gaaagtgccacactgacgtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2	AAV76919 Staphyloc
C 2	101	100.0	228	1	AAN10032 Sequence
C 3	101	100.0	251	1	AAN10031 Sequence
C 4	101	100.0	400	2	AAV31229 E. coli J
C 5	101	100.0	456	1	AAN60624 Plasmid p
C 6	101	100.0	456	1	AAN71080 Sequence
C 7	101	100.0	456	1	AAN70833 Beta-urog
C 8	101	100.0	456	1	AAN81765 Sequence
C 9	101	100.0	466	6	ABA90413 Drosophil
C 10	101	100.0	487	2	AAx21173 Polynucle
C 11	101	100.0	535	2	AAx21149 Polynucle
C 12	101	100.0	573	6	ABA90456 Drosophil
C 13	101	100.0	605	12	ADH58311 Electroph
C 14	101	100.0	776	4	AAS30560 DNA encod
C 15	101	100.0	776	4	AAS27819 DNA encod
C 16	101	100.0	776	4	ABK42984 Genomic s
C 17	101	100.0	776	4	AAL07344 Human rep
C 18	101	100.0	776	4	AAL03229 Human rep
C 19	101	100.0	776	4	AAL06588 Human rep
C 20	101	100.0	776	4	AAL07340 Human rep

C 21	101	100.0	776	5	ABA14573 Human ner
C 22	101	100.0	776	5	AAS34681 Human DNA
C 23	101	100.0	776	8	ADA41574 Human sec
C 24	101	100.0	776	8	ACC50905 Human sec
C 25	101	100.0	776	8	ABZ71508 Secreted
C 26	101	100.0	776	9	ADB91869 Human sec
C 27	101	100.0	776	9	ADB61140 Connectiv
C 28	101	100.0	776	10	ADB94622 Novel hum
C 29	101	100.0	776	10	ADC74663 Human sec
C 30	101	100.0	776	10	ADA57709 BAC fragm
C 31	101	100.0	776	12	ADN41551 Novel hum
C 32	101	100.0	845	4	AAS30559 DNA encod
C 33	101	100.0	845	4	AAS27818 DNA encod
C 34	101	100.0	845	4	ABK42983 Genomic s
C 35	101	100.0	845	4	AAS41807 Genomic s
C 36	101	100.0	845	4	AAS41855 Genomic s
C 37	101	100.0	845	4	AAK85485 Human imm
C 38	101	100.0	845	4	AAK85434 Human imm
C 39	101	100.0	845	4	AAL07343 Human rep
C 40	101	100.0	845	4	AAL06587 Human rep
C 41	101	100.0	845	4	AAL07339 Human rep
C 42	101	100.0	845	4	AAL03228 Human rep
C 43	101	100.0	845	5	ABA14572 Human ner
C 44	101	100.0	845	5	AAS34680 Human DNA
C 45	101	100.0	845	9	ADB61139 Connectiv

#### ALIGNMENTS

#### RESULT 1

AAV76919/c  
ID AAV76919 standard; DNA; 142 BP.  
XX AC AAV76919;  
XX DT 16-MAR-1999 (first entry)  
XX DB Staphylococcus aureus contig SEQ ID #2608.

KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.

OS Staphylococcus aureus.

PN EP786519-A2.

PD 30-JUL-1997.

PF 07-JAN-1997; 97EP-00100117.

PR 05-JAN-1996; 96US-0009861P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
XX stored on computer readable medium and used in the production of anti-  
XX S.aureus vaccines.

XX Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
XX of the invention. The DNA sequences are recorded on a computer readable  
XX medium, preferably selected from a floppy or hard disk, random access  
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
XX the S.aureus DNA sequences allows putative functions to be assigned so  
XX that protein-encoding or regulatory regions of commercial, therapeutic or



CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
CC fragment (CB6) for rat preproinsulin (see AAN10034)

XX Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60  
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 116  
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 101  
DB 115 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 75

## RESULT 4

AAV31229/c  
ID AAV31229 standard; DNA; 400 BP.

XX AC AAV31229;

XX DT 01-OCT-1998 (first entry)

XX DE E. coli J96 pathogenicity island contig #43.

XX KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;  
XX KW PAI V; pheV; vaccine; protective immune response; ds.

XX OS Escherichia coli.

XX PN WO9822575-A2.

XX PD 28-MAY-1998.

XX PF 21-NOV-1997; 97WO-US021347.

XX PR 22-NOV-1996; 96US-0031626P.

XX PR 14-OCT-1997; 97US-0061953P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (UYWI-) UNIV WISCONSIN.

XX PI Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to develop  
XX products for the detection of pathogenic E. coli and to elicit an immune  
XX response to pathogenic E. coli.

XX Claim 21; Page 140-141; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing  
XX pathogenicity island (PAI) sequences, and represents a nucleic acid  
XX molecule of the invention. PAIs are large fragments of DNA which comprise  
XX pathogenicity determinants. The sequences of the invention are taken from  
XX PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)  
XX on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
XX approximately 94 min (at pheR) on the E. coli chromosome and is  
XX approximately 160 kb in size. Antibodies specific to the proteins encoded  
XX by the PAI open reading frames of the invention can be used in kits to  
XX detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
XX a protective immune response in an animal to the uropathogenic E. coli  
XX strain J96

XX Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60  
DB 165 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 106  
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 101  
DB 105 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 65

## RESULT 5

AAN60624/c  
ID AAN60624 standard; DNA; 456 BP.

XX AC AAN60624;

XX DT 25-MAR-2003 (revised)

XX DT 29-OCT-1991 (first entry)

XX DE Plasmid pUG201 sequence encoding beta-urogastrone.

XX KW Beta-lactamase signal peptide; pGH54; pGH55; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT promoter 125..170

XX FT /\*tag= a

XX FT RBS 200..203

XX FT /\*tag= b

XX FT CDS 209..439

XX FT /\*tag= c

XX FT sig\_peptide 209..277

XX FT /\*tag= d

XX FT /label= Beta-lactamase signal peptide

XX FT 278..436

XX FT /\*tag= e

XX FT /label= Beta-urogastrone

XX WO8603779-A.

XX PD 03-JUL-1986.

XX PF 19-DEC-1985; 85WO-JP000696.

XX PR 21-DEC-1984; 84JP-00271206.

XX PA (EART ) EARTH CHEM CO LTD.

XX PA (OHGA/) OHGAI H.

XX PI Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

XX WPI; 1986-182911/28.

XX P-PSDB; AAP60678.

XX Recombinant vector for polypeptide secretion - contains signal peptide  
XX sequence directly bonded to peptide-coding sequence.

XX Disclosure; Table 4; 79pp; Japanese.

XX The plasmid produces secreted beta-urogastrone in a transformed  
XX expression system. Similar plasmids may be constructed where the  
XX secretion signal may be coupled with eg. somatostatin, insulin, growth  
XX hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,  
XX epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to  
XX correct PA field.)

XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 60  
|||||  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 114  
|||||

QY 61 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101  
|||||  
Db 113 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 73  
|||||

## RESULT 6

AAAN71080/c  
ID AAAN71080 standard; DNA; 456 BP.  
XX  
AC AAAN71080;  
XX  
DT 25-MAR-2003 (revised)  
DT 10-MAR-2003 (revised)  
DT 13-MAY-1991 (first entry)  
XX  
DE Sequence encoding beta-urogastrone.  
XX  
KW pUGT 150s; beta-UG; ds.  
XX  
XX Escherichia coli.  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT promoter 125..170  
FT /\*tag= a  
FT CDS 209..439  
FT /\*tag= b  
FT /\*transl\_except= (pos:434..436,aa:Arg)  
XX

JPG2190083-A.

XX  
XX 20-AUG-1987.  
XX  
XX 14-FEB-1986; 86JP-00031415.  
XX  
XX 14-FEB-1986; 86JP-00031415.  
XX  
XX (EART ) EARTH SEIYAKU KK.  
XX  
XX WPI; 1987-273761/39.  
XX  
XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.  
XX  
XX Disclosure; Page 553; 34pp; Japanese.  
XX  
XX Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing sequences comprising a tac promoter, SD site, signal peptide, and coding sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 60  
|||||  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 114  
|||||

QY 61 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101  
|||||  
Db 113 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 73  
|||||

## RESULT 7

AAAN70833/c  
ID AAAN70833 standard; DNA; 456 BP.  
XX  
AC AAAN70833;  
XX  
DT 25-MAR-2003 (revised)  
DT 10-MAR-2003 (revised)  
DT 18-JAN-1991 (first entry)  
XX  
DE Beta-urogastrone sequence.  
XX  
XX Tumour; inosine; DNA probe; ds.  
XX  
OS Unidentified.

FH Key Location/Qualifiers  
FT promoter 125..170  
FT /\*tag= b  
FT RBS 200..204  
FT /\*tag= c  
FT CDS 209..439  
FT /\*tag= a  
FT sig\_peptide 209..277  
FT /\*tag= d

JP62244398-A.

XX  
XX 24-OCT-1987.  
XX  
XX 16-APR-1986; 86JP-00087368.  
XX  
XX 16-APR-1986; 86JP-00087368.  
XX  
XX (SEKI ) SEKISUI CHEM IND CO LTD.  
XX  
XX WPI; 1987-339045/48.  
DR P-PSDB; AAP70505.

XX  
XX Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.

XX Disclosure; Page 11; 11pp; Japanese.

XX  
XX An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The ssDNA and probe are hybridized and the existence of DNA in the product is detected. It can be used to detect the presence of malignant tumour.  
XX  
XX (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
XX

SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 60  
|||||  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAATAG 114  
|||||

QY 61 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101  
|||||  
Db 113 GGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGACGTC 73  
|||||

## RESULT 8

AAAN81765/c  
ID AAAN81765 standard; DNA; 456 BP.  
XX  
XX AAAN81765;  
XX  
XX 25-MAR-2003 (revised)  
DT 13-DEC-1990 (first entry)

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),  
DE Arg (53).  
XX Gastric acid secretion; cell proliferation; hormone; ds.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH CDS 209..277 /\*tag= a  
FT CDS 278..439 /\*tag= b  
FT /product= "New beta-urogastrone deriv."  
XX JP63012298-A.  
XX 19-JAN-1988.  
XX 30-JUN-1986; 86JP-00153783.  
XX 30-JUN-1986; 86JP-00153783.  
XX (EART ) EARTH SEIYAKU KK.  
XX WPI; 1988-054638/08.  
XX P-PSDB; AAP81349.  
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and  
PT proliferation promotion activity.  
XX Disclosure; Page 685; 76pp; Japanese.  
XX The deriv. has various biological activities such as gastric acid  
CC secretion inhibiting action, or cell proliferation promoting action. The  
CC deriv. has the same biological or pharmacological activities as beta-  
CC urogastrone. It is not susceptible to denaturation by oxidn. and is  
CC chemically stable. Deriv. has resistance to proteolytic enzymes such as  
CC protease. (Updated on 25-MAR-2003 to correct PA field.)  
XX  
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114  
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
RESULT 9  
ABA90413/C  
ID ABA90413 standard; DNA; 466 BP.  
XX  
XX ABA90413;  
XX  
XX 12-FEB-2002 (first entry)  
XX Drosophila cell cycle progression protein coding sequence #48.  
DE Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;  
XX antiinflammatory; antiparasitic; dermatological; antifungal; mitosis;  
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;  
XX cell cycle progression protein; tumour; proliferative disorder;  
XX cardiovascular; autoimmune; dermatological disorder; ds.  
XX Drosophila sp.  
XX  
PN WO200172774-A2.  
XX 04-OCT-2001.  
XX 23-MAR-2001; 2001WO-GB001297.  
XX 24-MAR-2000; 2000GB-00007268.  
XX (CYCL-) CYCLACEL LTD.  
XX Deak P, Glover DM, Midgley C;  
XX WPI; 2002-055132/07.  
XX Polynucleotides encoding cell cycle progression proteins, useful for  
PT treating a tumor or a proliferative disorder.  
XX Claim 1; Page 99; 213pp; English.  
XX The present invention relates to Drosophila cell cycle progression  
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-  
CC ABA90520). The coding sequences and proteins are useful for identifying a  
CC substance capable of affecting the function of the corresponding gene, a  
CC substance capable of inhibiting the cell division cycle, or capable of  
CC inhibiting mitosis and/or meiosis. They can also be used in a method for  
CC treating a tumour or proliferative disorder, cardiovascular disorders  
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as  
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders  
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic  
CC disorders (such as malaria)  
XX  
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;  
Query Match 100.0%; Score 101; DB 6; Length 466;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60  
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 221  
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
Db 220 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180  
RESULT 10  
AAAX21173/C  
ID AAAX21173 standard; DNA; 487 BP.  
XX  
XX AAAX21173;  
XX  
XX 05-MAY-1999 (first entry)  
XX Polynucleotide sequence from the genome of Treponema pallidum.  
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;  
XX enzyme production; ds.  
XX Treponema pallidum.  
XX  
XX WO9859034-A2.  
XX  
XX 30-DEC-1998.  
XX 23-JUN-1998; 98WO-US013041.  
XX 24-JUN-1997; 97US-0050667P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Fraser CM;  
XX

DR WPI; 1999-081273/07.  
 XX New isolated *Treponema pallidum* nucleic acids - used to develop products  
 PT for the detection, diagnosis, characterisation, prevention and therapy of  
 PT *T. pallidum* infections, particularly syphilis.  
 XX  
 XX Claim 1; Page 1106; 1150pp; English.  
 XX  
 CC AAX20500-21243 represent polynucleotide sequences from the genome of  
 CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,  
 CC characterisation, prevention and therapy for *T. pallidum* infections,  
 CC particularly syphilis. They can also be used for detecting diseases  
 CC related to *Borrelia* infections in animals, and for the production of  
 CC biosynthetic products such as enzymes  
 XX  
 XX Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;  
 SQ

Query Match 100.0%; Score 101; DB 2; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX

Qy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTTGAATTTAGAAAAATAACAATAG 60  
 |||||  
 Db 323 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTTGAATTTAGAAAAATAACAATAG 264  
 |||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
 |||||  
 Db 263 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 223  
 |||||

RESULT 11  
 AAX21149/C  
 ID AAX21149 standard; DNA; 535 BP.  
 XX  
 AC AAX21149;  
 XX  
 XX 05-MAY-1999 (first entry)  
 DT  
 XX Polynucleotide sequence from the genome of *Treponema pallidum*.  
 DE  
 XX *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;  
 KW enzyme production; ds.  
 KW  
 OS *Treponema pallidum*.  
 OS  
 XX WO9859034-A2.  
 FN  
 XX 30-DEC-1998.  
 PD  
 XX 23-JUN-1998; 98WO-US013041.  
 PF  
 XX 24-JUN-1997; 97US-0050667P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Fraser CM;  
 PI  
 XX WPI; 1999-081273/07.  
 DR  
 XX New isolated *Treponema pallidum* nucleic acids - used to develop products  
 XX for the detection, diagnosis, characterisation, prevention and therapy of  
 XX *T. pallidum* infections, particularly syphilis.  
 XX  
 XX Claim 1; Page 1093; 1150pp; English.  
 PS  
 XX AAX20500-21243 represent polynucleotide sequences from the genome of  
 CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,  
 CC characterisation, prevention and therapy for *T. pallidum* infections,  
 CC particularly syphilis. They can also be used for detecting diseases  
 CC related to *Borrelia* infections in animals, and for the production of  
 CC biosynthetic products such as enzymes  
 XX  
 XX Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;  
 SQ

Query Match 100.0%; Score 101; DB 6; Length 573;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX

Qy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTTGAATTTAGAAAAATAACAATAG 60  
 |||||  
 Db 355 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTTGAATTTAGAAAAATAACAATAG 296  
 |||||

RESULT 12  
 ABA90456/C  
 ID ABA90456 standard; DNA; 573 BP.  
 XX  
 AC ABA90456;  
 XX  
 XX 12-FEB-2002 (first entry)  
 DT  
 XX *Drosophila* cell cycle progression protein coding sequence #91.  
 DE  
 XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;  
 KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;  
 KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;  
 KW cell cycle progression protein; tumour; proliferative disorder;  
 KW cardiovascular; autoimmune; dermatological disorder; ds.  
 XX  
 OS *Drosophila* sp.  
 OS  
 XX WO200172774-A2.  
 FN  
 XX 04-OCT-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-GB001297.  
 PF  
 XX 24-MAR-2000; 2000GB-00007268.  
 PR  
 XX (CYCL-) CYCLACEL LTD.  
 PA  
 XX Deak P, Glover DM, Midgley C;  
 PI  
 XX WPI; 2002-055132/07.  
 DR  
 XX Polynucleotides encoding cell cycle progression proteins, useful for  
 XX treating a tumor or a proliferative disorder.  
 PT  
 XX Claim 1; Page 144; 213pp; English.  
 PS  
 XX The present invention relates to *Drosophila* cell cycle progression  
 CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-  
 CC ABA90520). The coding sequences and proteins are useful for identifying a  
 CC substance capable of affecting the function of the corresponding gene, a  
 CC substance capable of inhibiting the cell division cycle, or capable of  
 CC inhibiting mitosis and/or meiosis. They can also be used in a method for  
 CC treating a tumour or proliferative disorder, cardiovascular disorders  
 CC (such as restenosis and cardiomyopathy), autoimmune disorders such as  
 CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders  
 CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic  
 CC disorders (such as malaria)  
 CC  
 XX Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;  
 SQ





PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234987P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249247P.  
PR 17-NOV-2000; 2000US-0249248P.  
PR 17-NOV-2000; 2000US-0249249P.  
PR 17-NOV-2000; 2000US-0249250P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250161P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251031P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251858P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251988P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-476223/51.

XX Novel isolated prostate gland related polypeptide useful for diagnosis  
PT and treatment of disorders of prostate such as prostatodystonia,  
PT prostatitis, prostatitis, benign prostatic hypertrophy and malacoplakia.

PS Claim 1; SEQ ID NO 418; 512pp; English.

XX The invention relates to novel isolated prostate gland related nucleic  
CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,  
CC prognosis, prevention, and/or treatment of diseases and/or disorders of  
CC the prostate such as acute non-bacterial prostatitis, chronic non-  
CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,  
CC prostatitis, granulomatous prostatitis, malacoplakia, benign prostatic  
CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including  
CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and  
CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for  
CC diagnosing and treating reproductive system disorders (Paget's disease),  
CC autoimmune disorders (systemic lupus erythematosus, rheumatoid  
CC arthritis), blood-related disorders (sickle cell anemia),  
CC hyperproliferative disorders, urinary system disorders  
CC (glomerulonephritis), cardiovascular system disorders (arrhythmias), respiratory  
CC disorders, musculoskeletal system disorders, neural activity and  
CC neurological disorders (Alzheimer's disease and Parkinson's disease),  
CC endocrine disorders (Addison's disease), gastrointestinal disorders  
CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),  
CC pancreatic and gall bladder disorders, disorders of the large intestine,  
CC developmental and inherited disorders, diseases at the cellular level,  
CC and wound healing and epithelial cell proliferation. (I) or (II) is  
CC useful to prevent skin aging, for preventing hair loss, to maintain  
CC organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;

Best Local Similarity 100.0%; Pred. No. 2.9e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTCCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAATAACAATAG 60

Db 546 AGGGTTATTCCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAATAACAATAG 487

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

Db 486 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 446

RESULT 15

AAS27819/c

ID AAS27819 standard; DNA; 776 BP.



PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251889P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465460/50.	
XX		
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or	
PT	prognosing disorders related to the proteins, including cancers, immune	
PT	disorders and neuronal disorders.	
XX		
PS	Claim 1; SEQ ID NO 1479; 880pp; English.	
XX		
CC	The invention relates to novel isolated polypeptides (I), and	
CC	polynucleotides (II). (I), (II) and the antibody to (I) are useful for	
CC	diagnosing, preventing and treating diseases including immune system	
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune	
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ	
CC	transplant rejections and graft versus host disease, infectious diseases	
CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and	
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative	
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders	
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities	
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.	
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),	
CC	respiratory disorders, dermatological disorders, in wound healing,	
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's	
CC	disease), reproductive system disorders, gastrointestinal disorder	
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of	
CC	B-cell responsiveness to pathogens, activators of T-cells, to induce	
CC	higher affinity antibodies, and as a means to induce tumour proliferation	
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-	
CC	AAS27850 represent novel signal transduction pathway protein coding	
CC	sequences and PCR primers of the invention	
XX		
SQ	Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;	
	Query Match 100.0%; Score 101; DB 4; Length 776;	
	Best Local Similarity 100.0%; Pred. No. 2.9e-21;	
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 AGGGTTATTGTCATGAGCGGTACATATTTGAATGATGATTTAGAAAAATAACAATAAG 60	
Dd	546 AGGGTTATTGTCATGAGCGGTACATATTTGAATGATGATTTAGAAAAATAACAATAAG 487	
QY	61 GGGTTCGCGCACATTTCCCCGAAAAGTGCCACCTGCAGTC 101	
Dd	486 GGGTTCGCGCACATTTCCCCGAAAAGTGCCACCTGCAGTC 446	

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-47\_COPY\_7889\_7989

Perfect score: 101

Sequence: 1 aggggtattgtctatgagc.....gaaagtgcacctgacgtc 101

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl1:\*

9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	AL000426
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819233
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

## ALIGNMENTS

RESULT 1

BM078095/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM078095 300 bp mRNA linear EST 30-NOV-2001  
83374 Hebeloma cylindrosporum functional cDNA library Hebeloma  
cylindrosporum cDNA 5', mRNA sequence.  
BM078095  
EST  
BM078095.1 GI:17157967  
Hebeloma cylindrosporum  
Hebeloma cylindrosporum  
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Agaricales; Cortinariaceae; Hebeloma.  
Wipf, D., Benjidia, M., Tegeder, M. and Frommer, W.B.  
Construction of a functional cDNA library from the ectomycorrhizal  
fungus Hebeloma cylindrosporum  
Unpublished (2001)  
Contact: Wipf D.  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: PDR196 5' primer (PWA 5')  
High quality sequence stop: 300  
POLYA=No.

FEATURES

source

Location/Qualifiers

1..300

/organism="Hebeloma cylindrosporum"

/mol\_type="mRNA"

/strain="H1"

/db\_xref="taxon:76867"

/tissue\_type="Mycelia"

/lab\_host="E. coli XLI-Blue"

/clone\_lib="Hebeloma cylindrosporum functional cDNA library"

/note="vector: pDR 196 (unpublished); Site\_1: EcoRI;

Site\_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 8.1e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGGTTATTGTCATGCGGACATACATATTGTAATGTTAGTATTAGAAAAATAACAAATAG 60

```

|||||
174 AGGGTTATTGCTCATGCGCGATACATATTGATGTTTGTAGAAAAATAACAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGATGTTTGTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGCGATACATATTGATGTTTGTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmpr.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
1 (bases 1 to 300)
Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGATGTTTGTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGCGATACATATTGATGTTTGTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmpr.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 5
AL597149
LOCUS
DEFINITION DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stages="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: sf1IA; Site_2: sf1IB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 6
AL597149
LOCUS
DEFINITION DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stages="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: sf1IA; Site_2: sf1IB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 6
AL597149
LOCUS
DEFINITION DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stages="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: sf1IA; Site_2: sf1IB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
mRNA sequence.
ACCESSION
VERSION BJ684174
KEYWORDS EST
SOURCE
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Labroidae; Cichlidae; Haplochromis.
REFERENCE
1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
Orf sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
JOURNAL
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
source
1..417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stage="varied"
/clone_lib="HCEST library"

ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
Db 129 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 70

Qy 61 GGGTTCGGCGACATTTCCCGAAAGTGCACCTGACGTC 101
Db 69 GGGTTCGGCGACATTTCCCGAAAGTGCACCTGACGTC 29

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10006J13 R, genomic survey
sequence.
ACCESSION
VERSION CC819923
KEYWORDS GSS.
SOURCE
ORGANISM Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotriches; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
Location/Qualifiers
source
1..491
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10006J13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/notes="Vector: FWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pMD42 (GI4732114|9b|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
Db 412 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 353

Qy 61 GGGTTCGGCGACATTTCCCGAAAGTGCACCTGACGTC 101
Db 352 GGGTTCGGCGACATTTCCCGAAAGTGCACCTGACGTC 312

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION
VERSION BI805285
KEYWORDS EST.
SOURCE
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
JOURNAL
COMMENT Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
Location/Qualifiers
source
1..495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

```



```

Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAAACAAATAG 60
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAAACAAATAG 332

QY 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
Db 332 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 292

RESULT 11
CC818523/c
LOCUS
DEFINITION
CC818523
CC818523.1 GI:32897943
GSS.
Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 496)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UTM,
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
Location/Qualifiers
1. 496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGCIO0004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGCIO library"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (GI_47321114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAAACAAATAG 60
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAAACAAATAG 332

QY 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

```

Db 331 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 291  
 RESULT 12  
 CC819854/c  
 LOCUS  
 DEFINITION  
 CC819854 503 bp DNA linear GSS 17-JUL-2003  
 100006N08R Oxytricha plasmid UUGC10 library Sterkiella  
 histriomuscorum genomic clone UUGC100006N08 R, genomic survey  
 sequence.  
 CC819854  
 CC819854.1 GI:32900533  
 GSS.  
 Sterkiella histriomuscorum (Oxytricha trifallax)  
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;  
 Stichotrichida; Oxytrichidae; Sterkiella.  
 1 (bases 1 to 503)  
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.  
 Paired end reads from plasmid inserts of Oxytricha trifallax  
 macronuclear chromosomes  
 Unpublished (2003)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Plate: 0006 row: N column: 08  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 503.  
 Location/Qualifiers  
 1..503  
 /organism="Sterkiella histriomuscorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:94289"  
 /clone="UUGC100006N08"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified macronuclear chromosomal  
 DNA from Oxytricha trifallax was blunt end-repaired with  
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
 oligonucleotides were ligated to the blunt ends in high  
 molar excess. Vector DNA was prepared from a derivative of  
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
 derivative of plasmid R1. The vector was ligated with  
 adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adapted vector DNA, and transformed into  
 chemically-competent E. Coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60  
 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 Db 410 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 351  
 QY 61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101  
 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 Db 350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310  
 ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; G

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 518)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0002 row: D column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 518.

FEATURES
    source
    1..518
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC100002D21"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /note="Vector: PWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adapted mouse DNA was annealed to
        adaptor vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

ORIGIN
Query Match          100.0%; Score 101; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAAATAACAATAG 60
    |||||||
Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAAATAACAATAG 351

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 350 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 310

RESULT 15
CC817162/c
LOCUS
DEFINITION
100002J199 Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100002J19 R, genomic survey
sequence.
ACCESSION
CC817162
VERSION
CC817162.1 GI:32896449
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 519)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center

```

```

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0002 row: J column: 19
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 519.

FEATURES
    source
    1..519
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC100002J19"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /note="Vector: PWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adapted mouse DNA was annealed to
        adaptor vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

ORIGIN
Query Match          100.0%; Score 101; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAAATAACAATAG 60
    |||||||
Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAAATAACAATAG 357

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 356 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:23:06
Job time : 962.667 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 ctgtctccctgtgtgtgtt.....caattgatgaagaattctgc 100

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.on.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	6	AX643583 Sequence
2	100	100.0	633	14	ALRPROL7B
3	100	100.0	648	6	AX175190 Sequence
4	100	100.0	648	6	AX175195 Sequence
5	100	100.0	1070	6	A85308 Sequence
6	100	100.0	1070	6	BD107647 FIV vacci
7	100	100.0	2245	6	AX643582 Sequence
8	100	100.0	2426	6	AX044426 Sequence
9	100	100.0	2427	6	AX044425 Sequence
10	100	100.0	3557	12	SYNRSV3MV
11	100	100.0	3840	12	EVE132038
12	100	100.0	3853	6	AR098190 Sequence
13	100	100.0	3853	6	AR207832 Sequence
14	100	100.0	3853	6	BD009729 Tissue sp
15	100	100.0	3925	6	A60213 Sequence
16	100	100.0	3925	6	AR122289 Sequence
17	100	100.0	3986	12	PCDNA32EO
18	100	100.0	4026	6	AR098191 Sequence
19	100	100.0	4026	6	AR207833 Sequence

c	20	100	100.0	4026	6	BD009730	BD009730 Tissue sp
	21	100	100.0	4059	6	AR071324	AR071324 Sequence
	22	100	100.0	4249	6	AR098192	AR098192 Sequence
	23	100	100.0	4249	6	AR207834	AR207834 Sequence
	24	100	100.0	4249	6	BD009731	BD009731 Tissue sp
	25	100	100.0	4341	6	A38214	A38214 Sequence 58
	26	100	100.0	4341	6	AX286570	AX286570 Sequence
	27	100	100.0	4457	6	AX743954	AX743954 Sequence
	28	100	100.0	4525	6	AR062871	AR062871 Sequence
	29	100	100.0	4597	6	AX060344	AX060344 Sequence
	30	100	100.0	4839	12	SYNRSV5GPT	M83236 Cloning vec
	31	100	100.0	4840	6	AX133940	AX133940 Sequence
	32	100	100.0	4965	6	AR071323	AR071323 Sequence
	33	100	100.0	5053	6	BD238492	BD238492 Expressio
	34	100	100.0	5070	6	AX234391	AX234391 Sequence
	35	100	100.0	5082	6	A91754	A91754 Sequence 10
	36	100	100.0	5082	6	BD085110	BD085110 Vertebrat
	37	100	100.0	5108	12	SYNRSV5NEO	M83237 cDNA expres
	38	100	100.0	5162	6	AX951626	AX951626 Sequence
	39	100	100.0	5257	12	CVU89673	U89673 Cloning vec
	40	100	100.0	5432	6	BD234590	BD234590 Screening
	41	100	100.0	5432	6	AX026821	AX026821 Sequence
	42	100	100.0	5446	6	BD195386	BD195386 Compositi
	43	100	100.0	5446	6	AX319694	AX319694 Sequence
	44	100	100.0	5564	12	SYNFCRC	L36555 cloning vec
	45	100	100.0	5618	6	A44171	A44171 Sequence 1

## ALIGNMENTS

RESULT 1  
AX643583  
LOCUS AX643583 562 bp DNA linear PAT 24-FEB-2003  
DEFINITION Sequence 2 from Patent WO02099100.  
ACCESSION AX643583  
VERSION AX643583.1 GI:28551383  
KEYWORDS  
SOURCE Mus 'sp'.  
ORGANISM Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Al-Rubeai, M. and Shuttleworth, J.  
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21  
JOURNAL Patent: WO 02099100-A 2 12-DEC-2002;  
FEATURES Lonza Biologics plc (GB)  
Location/Qualifiers  
source 1..562  
/organism="Mus sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10095"  
/note="Rous Sarcoma Virus LTR promoter"

ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 562;  
Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTAAGCTACA 60  
Db 46 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTAAGCTACA 105  
Qy 61 ACAAGCAAGCTTGCACCGCAATTCATGAAGAAATCTGC 100  
Db 106 ACAAGCAAGCTTGCACCGCAATTCATGAAGAAATCTGC 145

RESULT 2  
ALRPROL7B  
LOCUS ALRPROL7B 633 bp ss-RNA linear VRL 28-APR-1993  
DEFINITION Rous sarcoma virus (Schmidt-Ruppin), proviral, 3' LTR on 21S mRNA.

J02025 J02022  
 J02025.1 GI:210255  
 C-myc proto-oncogene; long terminal repeat (LTR); src oncogene.  
 Rous sarcoma virus  
 Rous sarcoma virus  
 Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.  
 1 (sites)  
 Yamamoto,T., de Crombrughe,B. and Pastan,I.  
 Identification of a functional promoter in the long terminal repeat  
 of Rous sarcoma virus  
 Cell 22 (3), 787-797 (1980)  
 81112147  
 PUBLISHED  
 6257399  
 2 (bases 1 to 633)  
 Yamamoto,T., Tyagi,J.S., Pagan,J.B., Jay,G., deCrombrughe,B. and  
 Pastan,I.  
 Molecular mechanism for the capture and excision of the  
 transforming gene of avian sarcoma virus as suggested by analysis  
 of recombinant clones  
 J. Virol. 35 (2), 436-443 (1980)  
 81072438  
 PUBLISHED  
 6255184  
 3 (bases 319 to 633)  
 Yamamoto,T., Jay,G. and Pastan,I.  
 Unusual features in the nucleotide sequence of a cDNA clone derived  
 from the common region of avian sarcoma virus messenger RNA  
 Proc. Natl. Acad. Sci. U.S.A. 77 (1), 176-180 (1980)  
 80145590  
 PUBLISHED  
 6244542  
 Original source text: Rous sarcoma virus (Schmidt-Ruppin strain,  
 subgroup D) provirus, cDNA to 21S mRNA from infected chicken  
 embryonic fibroblasts, clone pSR1.  
 [1] sites; mRNA start.  
 Original figure in [2] included 24 'g's on 5' end and 16 'c's on 3'  
 end that were cDNA synthesis artifacts.  
 [2] also sequenced a defective clone, pSR2, with the src gene  
 deleted (see separate entry).  
 [1] demonstrated the mRNA transcription initiation site shown in  
 the Sites table using pSR1 as a template. However, this is the 3',  
 LTR, and the functional mRNA start site would be assumed to be on  
 the 5' LTR at the homologous site.  
 FEATURES Location/Qualifiers  
 source 1..633  
 /organism="Rous sarcoma virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:11886"  
 misc\_RNA <1..517  
 /note="viroion genomic RNA"  
 LTR 211..>633  
 /note="3' LTR"  
 mRNA 517..>633  
 /note="in vitro mRNA [1]; see comment"  
 repeat\_region 517..536  
 /note="terminally redundant repeat"  
 ORIGIN 20 bp upstream of pET1 site.  
 Query Match 100.0%; Score 100; DB 14; Length 633;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
 |||||||  
 Db 28 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 87  
 |||||||  
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
 |||||||  
 Db 88 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 127  
 |||||||  
 RESULT 3  
 AX175190 648 bp DNA linear PAT 03-JUL-2001  
 LOCUS Sequence 1 from Patent WO014244.  
 DEFINITION

AX175190  
 AX175190.1 GI:14598581  
 synthetic construct  
 synthetic construct  
 other sequences; artificial sequences.  
 1  
 Rivera,V., Zoltick,P. and Wilson,J.M.  
 Methods for expression of genes in primates  
 Patent: WO 014244-A 1 14-JUN-2001;  
 ARIAD GENE THERAPEUTICS, INC. (US) ; THE UNIVERSITY OF PENNSYLVANIA  
 (US)  
 FEATURES Location/Qualifiers  
 source 1..648  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="vector/RSV promoter/vector"  
 ORIGIN  
 Query Match 100.0%; Score 100; DB 6; Length 648;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
 |||||||  
 Db 90 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 149  
 |||||||  
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
 |||||||  
 Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189  
 |||||||  
 RESULT 4  
 AX175195 648 bp DNA linear PAT 03-JUL-2001  
 LOCUS Sequence 6 from Patent WO014244.  
 AX175195  
 AX175195  
 AX175195.1 GI:14598586  
 synthetic construct  
 synthetic construct  
 other sequences; artificial sequences.  
 1  
 Rivera,V., Zoltick,P. and Wilson,J.M.  
 Methods for expression of genes in primates  
 Patent: WO 014244-A 6 14-JUN-2001;  
 ARIAD GENE THERAPEUTICS, INC. (US) ; THE UNIVERSITY OF PENNSYLVANIA  
 (US)  
 FEATURES Location/Qualifiers  
 source 1..648  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="MluI/RSV promoter/BglI"  
 ORIGIN  
 Query Match 100.0%; Score 100; DB 6; Length 648;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
 |||||||  
 Db 90 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 149  
 |||||||  
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
 |||||||  
 Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189  
 |||||||  
 RESULT 5  
 A85308 1070 bp DNA linear PAT 21-JAN-2000  
 LOCUS

```

DEFINITION Sequence 6 from Patent WO9840493.
ACCESSION A85308
VERSION A85308.1 GI:6733916
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Rigby,M.A. and Jarrett,J.O.
TITLE FIV VACCINE
JOURNAL Patent: WO 9840493-A 6 17-SEP-1998;
RIGBY MARK ALAN (GB); JARRETT JAMES OSWALD (GB)
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176
RESULT 7
AX0443582
LOCUS AX0443582 2245 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 1 from Patent WO02099100.
ACCESSION AX0443582
VERSION AX0443582.1 GI:28551382
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1
AUTHORS Al-Rubeai,M. and Shuttleworth,J.
TITLE Method of production of a protein in cells which inducibly express
the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 1 12-DEC-2002;
Lonza Biologics plc (GB)
FEATURES
    source
        Location/Qualifiers
            1..2245
                /organism="Mus sp."
                /mol_type="unassigned DNA"
                /db_xref="taxon:10095"
                /note="RSV-LTR promoter + intron + p21 cds + Tkpoly(A)
                LacSwitch II expression construct"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 2245;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 46 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 105
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145
RESULT 8
AX044426
LOCUS AX044426 2426 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 18 from Patent WO0066752.
ACCESSION AX044426
VERSION AX044426.1 GI:11343299
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Castro,M.G., Emery,S.C. and Lowenstein,P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 18 09-NOV-2000;
Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
    source
        Location/Qualifiers
            1..2426
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="CPG2 with last exon of Thy-1 fused at 3' end"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176
RESULT 6
BD107647
LOCUS BD107647 1070 bp DNA linear PAT 18-SEP-2002
DEFINITION FIV vaccine.
ACCESSION BD107647
VERSION BD107647.1 GI:23202465
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Neil,J.C., Rigby,M.A. and Jarrett,J.O.
TITLE FIV vaccine
JOURNAL Patent: JP 2002501369-A 6 15-JAN-2002;
THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW
COMMENT OS Artificial Sequence
OS Unknown
PN JP 2002501369-A/6
PD 15-JAN-2002
PF 10-MAR-1998 JP 1998539351
PR 11-MAR-1997 GB 9704977.9
PI JAMES CHARLES NEIL,MARK ALAN RIGBY,JAMES OSWALD JARRETT PC
C12N15/49,A61K31/70,A61K48/00
CC CMV PROMOTER FROM pCDNA3(a Bgl II - Kpn
I restriction fragment)
CC SST I - SST I FRAGMENT IN PLASMID CMV DEL. RT CC FIV GENOME
FROM THE t-RNA PRIMER BINDING
SITE TO THE VIRAL SAT
CC IS
FH Key Location/Qualifiers
FT source 1..1070
    /organism='Artificial Sequence'.
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

```

Query Match 100.0%; Score 100; DB 6; Length 2426;  
Best Local Similarity 100.0%; Pred. No. 3.1e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
|||||  
Db 69 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 128  
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 129 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 168  
|||||

RESULT 9  
LOCUS AX044425 2427 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 17 from Patent WO0066752.  
ACCESSION AX044425  
VERSION AX044425.1 GI:11343298  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.  
TITLE Chemical compounds  
JOURNAL Patent: WO 0066752-A 17 09-NOV-2000;  
Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES  
source  
1. 2427  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CPG2 mutant with last exon of Thy-1 fused at 3'  
end"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 2427;  
Best Local Similarity 100.0%; Pred. No. 3.1e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
|||||  
Db 70 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 129  
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 130 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 169  
|||||

RESULT 10  
LOCUS SYNRSV3MV 3557 bp DNA circular SYN 27-APR-1993  
DEFINITION Cloning vector RSV3.  
ACCESSION M83240  
VERSION M83240.1 GI:209303  
KEYWORDS cDNA expression vector.  
SOURCE unidentified cloning vector  
ORGANISM other sequences; artificial sequences; vectors.

REFERENCE 1  
AUTHORS Messing, J.  
TITLE New M13 vectors for cloning  
JOURNAL Meth. Enzymol. 101, 20-78 (1983)  
MEDLINE 83296918  
PUBMED 6310323

REFERENCE 2  
AUTHORS Gorman, C., Padmanabhan, R. and Howard, B.H.  
TITLE High efficiency DNA-mediated transformation of primate cells  
JOURNAL Science 221 (4610), 551-553 (1983)  
MEDLINE 83249156  
PUBMED 6306768

REFERENCE 3  
AUTHORS (bases 1 to 3557)

AUTHORS Jacobson, S., Sekaly, R.P., Jacobson, C.L., McFarland, H.F. and Long, B.O.  
TITLE HUA class II-restricted presentation of cytoplasmic measles virus antigens to cytotoxic T cells  
JOURNAL J. Virol. 63 (4), 1756-1762 (1989)  
MEDLINE 89178863  
PUBMED 2784508  
COMMENT Original source text: Cloning vector DNA.  
FEATURES  
source  
1. 3557  
Location/Qualifiers  
/organism="unidentified cloning vector"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:45196"  
misc\_feature  
1. 29  
/function="polylinker"  
/evidence="experimental"  
912..3029  
misc\_feature  
/function="ampicillin-resistance, replication origin"  
/evidence="experimental"  
3030..3557  
enhancer  
/standard\_name="5'LTR of Rous Sarcoma Virus"  
/citation=[2]  
/evidence="experimental"

ORIGIN

Query Match 100.0%; Score 100; DB 12; Length 3557;  
Best Local Similarity 100.0%; Pred. No. 3.2e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
|||||  
Db 3030 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 3089  
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 3090 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 3129  
|||||

RESULT 11  
LOCUS EVE132038 3840 bp RNA circular SYN 28-JUL-1999  
DEFINITION Expression vector pCDPT.  
ACCESSION AJ132038  
VERSION AJ132038.1 GI:5640088  
KEYWORDS AMP gene; beta lactamase; Cole1 origin of replication; multiple cloning site; SP6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; xanthine-guanine phosphoribosyl transferase gene.

SOURCE  
ORGANISM Expression vector pCDPT  
other sequences; artificial sequences; vectors.

REFERENCE 1  
AUTHORS Zeng, B.J.  
TITLE Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase tag  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3840)  
AUTHORS Zeng, B.J.  
TITLE Direct Submission  
JOURNAL Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA

FEATURES  
source  
1. 3840  
Location/Qualifiers  
/organism="Expression vector pCDPT"  
/mol\_type="other RNA"  
/db\_xref="taxon:90749"  
209..863  
promoter  
/note="CMV"  
864..882  
promoter  
/note="T7"  
882..984  
misc\_feature  
/note="Multiple cloning site"



```

CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSKSYIYTWMLQIHARKLASRLMPSEQWKGIIIVSRGGILVPCA
LLARELGHVHDVTVCISYDHNQREILKVLKRAEGDGEFIVDDLVDTGGTAVAIHE
MYPKAFVTTIPAKPAGRLVDDYVDYIPQDTWIRQPDGMGVFVPPISGR"
1649..1863
/feature="BGN"
2450..2775
/feature="SP6"
2644..2729
/feature="SV40"
complement(2844..3704)
/feature="amp"
complement(2844..3704)
/feature="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHPRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSKILESFRPEFPMSTFKLLGAVLSRIDAGQEQLRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNNGDHVTSL
DRWPELNEAIIPNDRITTPVAMATTLKLLTGLLTLLASRQQLIDWMEADKVGPL
LRSLPAGWFTADKSGAGERSGIIAALGPDGKPSRIVVIYTTGSGQTWDERNRQIA
EIGASLIKHW"
3632..3840
/feature="ColEI"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
Db 6 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 65

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
LOCATION/Qualifiers
FEATURES
source
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 140

polyA_site
promoter
rep_origin
gene
CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSKSYIYTWMLQIHARKLASRLMPSEQWKGIIIVSRGGILVPCA
LLARELGHVHDVTVCISYDHNQREILKVLKRAEGDGEFIVDDLVDTGGTAVAIHE
MYPKAFVTTIPAKPAGRLVDDYVDYIPQDTWIRQPDGMGVFVPPISGR"
1649..1863
/feature="BGN"
2450..2775
/feature="SP6"
2644..2729
/feature="SV40"
complement(2844..3704)
/feature="amp"
complement(2844..3704)
/feature="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHPRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSKILESFRPEFPMSTFKLLGAVLSRIDAGQEQLRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNNGDHVTSL
DRWPELNEAIIPNDRITTPVAMATTLKLLTGLLTLLASRQQLIDWMEADKVGPL
LRSLPAGWFTADKSGAGERSGIIAALGPDGKPSRIVVIYTTGSGQTWDERNRQIA
EIGASLIKHW"
3632..3840
/feature="ColEI"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
Db 6 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 65

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
LOCATION/Qualifiers
FEATURES
source
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 140

```

```

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES
source
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 ctgtccctgtgtgtgtt.....caattgcatgaagaattgc 100

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2002bs.\*
- 8: Geneseq2003as.\*
- 9: Geneseq2003bs.\*
- 10: Geneseq2003cs.\*
- 11: Geneseq2003ds.\*
- 12: Geneseq2004as.\*
- 13: Geneseq2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	8	ABZ23250 Nucleotid
2	100	100.0	648	4	AH43951 Rous sarc
3	100	100.0	1070	2	AAV58058 Plasmid C
4	100	100.0	1506	12	ADMA1035 Fungus nu
5	100	100.0	1600	2	ADH11349 Vertebrat
6	100	100.0	1782	12	ADMA1037 Cytomegal
7	100	100.0	2241	12	ADMA1034 Human nuc
8	100	100.0	2245	8	ABZ23249 Lac repre
9	100	100.0	2294	12	ADMA1036 Cytomegal
10	100	100.0	2426	4	AD02037 Plasmid p
11	100	100.0	2427	4	AD02036 Plasmid p
12	100	100.0	3400	2	AAT62937 3F4 human
13	100	100.0	3400	2	AAT62932 2A2 human
14	100	100.0	3853	2	AAV40006 Plasmid p
15	100	100.0	3925	2	AAT90695 Plasmid C
16	100	100.0	4026	2	AAV40007 Plasmid p
17	100	100.0	4059	2	AAQ75974 pHLA-B7 e
18	100	100.0	4249	2	AAV63466 Plasmid p
19	100	100.0	4341	2	AAQ62391 Vector pV
20	100	100.0	4341	6	AAAS17704 Vector pV

21	100	100.0	4341	6	ABN83143	Abn83143 Plasmid p
22	100	100.0	4457	10	ADD35599	Add35599 Bicistron
23	100	100.0	4525	2	AAV69746	Aav69746 Nucleotid
24	100	100.0	4597	4	AAF24901	Aaf24901 Nucleotid
25	100	100.0	4825	13	ADR12380	Adr12380 Vector pM
26	100	100.0	4840	4	AAF83146	Aaf83146 Complete
27	100	100.0	4965	2	AAQ75973	Aaq75973 pHLA-B7/b
28	100	100.0	5015	10	ADB33528	Adb33528 Expressio
29	100	100.0	5053	3	AAZ38633	Aaz38633 pEP2 expr
30	100	100.0	5070	4	AAAS12839	Aas12839 DNA sequ
31	100	100.0	5082	2	ADH11417	Adh11417 plasmid p
32	100	100.0	5162	10	ADF10526	Adf10526 Plasmid p
33	100	100.0	5162	10	ACC44637	Acc44637 Murine rD
34	100	100.0	5172	13	ADS75099	Ads75099 Plasmid p
35	100	100.0	5173	6	ABK88869	Abk88869 Topoisome
36	100	100.0	5173	12	ADE83792	Ades83792 Plasmid p
37	100	100.0	5173	12	ADO06721	Ado06721 Recombina
38	100	100.0	5192	10	ACC44692	Acc44692 Plasmid p
39	100	100.0	5250	2	AAT62933	Aat62933 2A2 human
40	100	100.0	5271	10	ABV77540	Abv77540 Plasmid p
41	100	100.0	5283	10	ABV77538	Abv77538 Plasmid p
42	100	100.0	5292	10	ABV77547	Abv77547 Plasmid p
43	100	100.0	5293	10	ABV77548	Abv77548 Plasmid p
44	100	100.0	5293	10	ABV77549	Abv77549 Plasmid p
45	100	100.0	5300	2	AAT62938	Aat62938 3F4 human

ALIGNMENTS

RESULT 1  
ABZ23250  
ID ABZ23250 standard; DNA; 562 BP.  
AC ABZ23250;  
XX  
DT 24-MAR-2003 (first entry)  
XX  
DB Nucleotide sequence of the Rous sarcoma virus (RSV)-LTR promoter.  
KW p21; RSV; LTR promoter; cell cycle inhibitor protein; protein production;  
KW anchorage-independent producer cell line; ss.  
XX  
OS Rous sarcoma virus.  
XX  
PN WO200299100-A2.  
XX  
PD 12-DEC-2002.  
XX  
PF 03-JUN-2002; 2002WO-EP006054.  
XX  
PR 01-JUN-2001; 2001GB-00013318.  
XX  
PA (LONZ ) LONZA BIOLOGICS PLC.  
PI Al-Rubeai M, Shuttleworth J;  
XX  
DR WPI; 2003-148669/14.  
XX  
PT Producing recombinant protein, particularly for maximizing or enhancing  
PT e.g. therapeutic protein production, by co-expressing protein with  
PT recombinant cell cycle inhibitor protein (p21) in producer cell line.  
XX  
PS Disclosure; Page 32-33; 33pp; English.  
XX  
CC The present sequence represents the Rous sarcoma virus (RSV)-LTR  
CC promoter. The present sequence is used to produce vectors for use in the  
CC method of the invention. The specification describes a method for  
CC producing a protein, preferably a recombinant protein, in a mammalian  
CC anchorage-independent producer cell line. The method comprises co-  
CC expressing with the protein in the producer cell line a recombinant cell  
CC cycle inhibitor protein (preferably p21). The method is useful for  
CC producing a recombinant protein in a producer cell line. This is

CC particularly useful for maximizing or enhancing the production of e.g.  
XX therapeutic proteins at an industrial scale  
SQ Sequence 562 BP; 143 A; 109 C; 163 G; 147 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 8; Length 562;  
Best Local Similarity 100.0%; Pred. No. 8.2e-28;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
Db 46 CTGCTCCCTGCTTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 105  
Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100  
Db 106 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 145  
RESULT 2  
AAH43951  
ID AAH43951 standard; DNA; 648 BP.  
XX  
AC AAH43951;  
XX  
DT 06-SEP-2001 (first entry)  
XX  
DE Rous sarcoma virus promoter nucleotide sequence SEQ ID NO:1.  
XX  
XX Rous sarcoma virus; promoter; enhancer; RSV; primate; gene expression;  
KW transgene; genetic engineering; gene therapy; immunisation; ds.  
XX  
XX Rous sarcoma virus.  
OS  
XX WO200142444-A2.  
FN  
XX  
PD 14-JUN-2001.  
XX  
XX  
PF 08-DEC-2000; 2000WO-US033256.  
XX  
PR 10-DEC-1999; 99US-0170019P.  
XX  
PA (ARIA-) ARIAD GENE THERAPEUTICS INC.  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX  
Pi Rivera V, Zoltick P, Wilson JM;  
XX  
XX WPI; 2001-381673/40.  
XX  
XX Genetically engineering a primate for expression of a desired gene,  
PT comprises introducing into the primate a transgene comprising Rous  
PT Sarcoma Virus (RSV) promoter and a nucleic acid sequence heterologous to  
PT RSV promoter.  
XX  
XX Claim 7; Page 44; 64pp; English.  
XX  
XX  
XX The present invention describes a method for genetically engineering a  
CC primate for expression of a desired gene comprising introducing into the  
CC primate a transgene comprising an Rous Sarcoma Virus (RSV) promoter and a  
CC nucleic acid sequence heterologous to RSV promoter. Also described is a  
CC primate cell (I) containing and capable of expressing a transgene  
CC comprising an RSV promoter operably linked to a recombinant nucleic acid  
CC encoding one or more fusion proteins, where the fusion proteins bind to a  
CC ligand and in the presence of the ligand modulate(s) the expression level  
CC of a target gene. The method can be used for high level expression of  
CC genes in primates or for engineering primate cells. It is useful for  
CC increasing the efficacy of many gene therapy strategies, and for  
CC increasing the efficacy of intracellular immunisation agents, molecules  
CC like ribozymes, antisense RNA, and dominant negative proteins, that act  
CC either stoichiometrically, or by competition. The method increases the  
CC efficacy of many gene therapy strategies by substantially elevating the  
CC expression of an exogenous therapeutic gene, and allowing expression to  
CC reach therapeutically effective levels. The present sequence represents a  
CC specifically claimed RSV enhancer/promoter nucleotide sequence from the

CC present invention  
XX  
SQ Sequence 648 BP; 163 A; 135 C; 179 G; 171 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 4; Length 648;  
Best Local Similarity 100.0%; Pred. No. 8.6e-28;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60  
Db 90 CTGCTCCCTGCTTGTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 149  
Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100  
Db 150 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 189  
RESULT 3  
AAV58058  
ID AAV58058 standard; DNA; 1070 BP.  
XX  
AC AAV58058;  
XX  
DT 27-AUG-2003 (revised)  
DT 11-JAN-1999 (first entry)  
XX  
XX Plasmid CMV-delRT SstI fragment.  
DE  
XX  
XX FIV; FIPV; vaccine; reverse transcriptase; diagnosis; therapy; CMV-delRT;  
KW promoter; cat; ss.  
XX  
XX Human cytomegalovirus.  
OS feline immunodeficiency virus.  
OS Chimeric.  
XX  
XX  
XX Key Location/Qualifiers  
FT Promoter 8...896  
FT /\*tag= a  
FT /note= "CMV promoter fragment from pcDNA3 (BgIII-KpnII)"  
FT 918..1070  
FT /\*tag= b  
FT /note= "FIV sequences from primer binding site to SstI  
FT site"  
XX  
XX WO9840493-A1.  
PN  
XX  
XX 17-SEP-1998.  
PD  
XX  
XX 10-MAR-1998; 98WO-GB000715.  
PF  
XX  
XX 11-MAR-1997; 97GB-00004977.  
PR  
XX  
XX (UNIU ) UNIV GLASGOW.  
PA  
XX  
XX Neil JC, Rigby MA, Jarrett JO;  
PI  
XX  
XX WPI; 1998-520813/44.  
DR  
XX  
XX  
XX Protecting, e.g. cats, against feline immunodeficiency virus - by using  
PT vaccine comprising FIV pol gene containing deletion and/or insertion in  
PT reverse transcriptase domain.  
XX  
XX Example 3; Fig 4; 66pp; English.  
PS  
XX  
XX This is the nucleotide sequence of a SstI fragment of plasmid CMV-delRT,  
CC in which the immediate-early promoter of human cytomegalovirus replaces  
CC the 5' long terminal repeat region of feline immunodeficiency virus (FIV)  
CC clone F14-delRT (see AAV58053). FIV sequences downstream of the SstI site  
CC are identical to those in F14-delRT. Use of the CMV promoter was designed  
CC to enhance expression of FIV antigens, and to reduce the risk of  
CC reversion to a replicating provirus, in tissues after inoculation of DNA.  
CC Vaccine formulations for FIV-related diseases include a defective feline  
CC immunodeficiency proviral (FIPV) polynucleotide comprising an in-frame

CC deletion and/or insertion in the reverse transcriptase (RT) region of the  
CC pol gene. Host cells comprising the PIPV are capable of producing FIV  
CC proteins, except for functionally competent RT, and thus release non-  
CC infectious FIV viral particles. (Updated on 27-AUG-2003 to correct OS  
CC field.)

XX Sequence 1070 BP; 275 A; 254 C; 268 G; 273 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1070;

Best Local Similarity 100.0%; Pred. No. 1e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
Db 77 CTGCTCCCTGCTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 136

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAGATCTGC 100

Db 137 ACAAGGCAAGCTTGACCGACAATTGCATGAGATCTGC 176

#### RESULT 4

ADM41035

ID ADM41035 standard; DNA; 1506 BP.

XX

AC ADM41035;

XX

DT 17-JUN-2004 (first entry)

XX

DE Fungus nucleotide sequence SEQ ID NO:3.

XX

XX engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; fungus; gene; ds.

XX

OS Unidentified.

XX

PN WO2004027029-A2.

XX

PD 01-APR-2004.

XX

PF 17-SEP-2003; 2003WO-US029251.

XX

PR 19-SEP-2002; 2002US-0411790P.

XX

PA (XIME-) XIMEREX INC.

XX

PI Beschorner WE, Sosa CE, Thompson SC;

XX

DR WPI; 2004-295402/27.

XX

PS Engrafting foreign replacement cells within a fetal non-human mammal,

XX

CC useful in producing chimeric mammals, comprises selectively destroying

CC native cells in a tissue of a fetal non-human mammal host.

XX

XX Disclosure; SEQ ID NO 3; 48pp; English.

XX

CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises

CC selectively destroying native cells in a tissue of a foetal non-human

CC mammal host, where the number of maternal cells of the same tissue is not

CC substantially reduced, and implanting foreign replacement cells in the

CC tissue of the fetal non-human mammal host, where the foreign replacement

CC cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to

CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present

CC sequence represents a nucleotide sequence given in the Sequence Listing

CC of the present invention but not mentioned further within the

XX specification.

XX Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
Db 81 CTGCTCCCTGCTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAGATCTGC 100

Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAGATCTGC 180

#### RESULT 5

ADH11349

ID ADH11349 standard; DNA; 1600 BP.

XX

AC ADH11349;

XX

DT 11-MAR-2004 (first entry)

XX

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX

UN-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;

KW cell shape regulator; cell motility regulator; cell migration;

KW cell behaviour regulator; phenotype; signal transduction pathway;

KW signal transducing protein; signal integrator protein;

KW neuronal regeneration; revascularisation; wound healing;

KW chronic neurodegenerative disease; acute traumatic injury;

KW fibrotic disease; gene; ds.

XX

OS Unidentified.

XX

PN WO9824810-A2.

XX

PD 11-JUN-1998.

XX

PF 03-DEC-1997; 97WO-EP006956.

XX

PR 04-DEC-1996; 96GB-00025283.

XX

PA (JANC ) JANSSEN PHARM NV.

XX

PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

PI Geysen J, Bogaert TA0E;

XX

DR WPI; 1998-362411/31.

XX

DR P-PSDB; ADH11350.

XX

PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

PT promoting neuronal regeneration, treating chronic neuro-degenerative

PT diseases or acute traumatic injuries.

XX

XX Disclosure; Page 410-411; 479pp; English.

XX

CC The present invention describes a vertebrate protein homologue of an UNC-

CC 53 protein of Caenorhabditis elegans or a functional equivalent,

CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence

CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a

CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising

CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)

CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of

CC cell shape, motility, or the direction of cell migration for use as a

CC therapeutic; (7) a method for determination of whether a protein is an

CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or

CC motility or the direction of migration by contacting a host cell

CC expressing a homologue of UNC-53 and determining a change of phenotype;

CC (8) a method for identification of vertebrate homologues of C. elegans

CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to

CC a DNA library; and (9) a method for identification of a protein which is

CC active in the signal transduction pathway of a cell of which a vertebrate  
CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
CC antibody/homologue complex; and (iii) analysing such a complex to  
CC identify any non-antibody protein bound to the complex. UNC-53 is a  
CC signal transducing or signal integrator protein involved in controlling  
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate  
CC homologues of UNC-53 can be used to promote neuronal regeneration,  
CC revascularisation or wound healing, to treat chronic neurodegenerative  
CC diseases or acute traumatic injuries or fibrotic diseases. The present  
CC sequence is used in the exemplification of the present invention.

XX  
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGGAGGTGCTGAGTAGTGGCGAGCGAGCAAAATTTAAGCTACA 60  
DB 81 CTGCTCCCTGCTGTGTGTGGAGGTGCTGAGTAGTGGCGAGCGAGCAAAATTTAAGCTACA 140

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 6  
ADM41037  
ID ADM41037 standard; DNA; 1782 BP.  
AC ADM41037;  
XX  
XX  
DT 17-JUN-2004 (first entry)  
XX  
XX  
DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.  
XX  
XX  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
XX  
XX  
OS Cytomegalovirus.  
XX  
XX  
FN WO2004027029-A2.  
XX  
XX  
PD 01-APR-2004.  
XX  
XX  
PF 17-SEP-2003; 2003WO-US029251.  
XX  
XX  
PR 19-SEP-2002; 2002US-0411790P.  
XX  
XX  
PA (XIME-) XIMEREX INC.  
XX  
XX  
PI Beschorner WE, Sosa CE, Thompson SC;  
XX  
XX  
DR WPI; 2004-295402/27.  
XX  
XX  
PT Engrafting foreign replacement cells within a fetal non-human mammal,  
PT useful in producing chimeric mammals, comprises selectively destroying  
PT native cells in a tissue of a fetal non-human mammal host.  
XX  
XX  
PS Disclosure; SEQ ID NO 5; 48pp; English.

XX The present invention describes a method for engrafting foreign  
XX replacement cells within a fetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a fetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the fetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX  
SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 12; Length 1782;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGTGGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
DB 81 CTGCTCCCTGCTGTGTGTGGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 140

QY 61 ACAAGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
DB 141 ACAAGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 7  
ADM41034  
ID ADM41034 standard; DNA; 2241 BP.  
AC ADM41034;  
XX  
XX  
DT 17-JUN-2004 (first entry)  
XX  
XX  
DE Human nucleotide sequence SEQ ID NO:2.  
XX  
XX  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; human; gene; ds.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
FN WO2004027029-A2.  
XX  
XX  
PD 01-APR-2004.  
XX  
XX  
PF 17-SEP-2003; 2003WO-US029251.  
XX  
XX  
PR 19-SEP-2002; 2002US-0411790P.  
XX  
XX  
PA (XIME-) XIMEREX INC.  
XX  
XX  
PI Beschorner WE, Sosa CE, Thompson SC;  
XX  
XX  
DR WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a fetal non-human mammal host.  
XX  
XX  
PS Disclosure; SEQ ID NO 2; 48pp; English.

XX The present invention describes a method for engrafting foreign  
XX replacement cells within a fetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a fetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the fetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for  
XX transplantation, also useful to study human diseases. The present  
XX sequence represents a nucleotide sequence given in the Sequence Listing  
XX of the present invention but not mentioned further within the  
XX specification.

XX Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;  
Best Local Similarity 100.0%; Pred. No. 1.3e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTCTGCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 60  
Db 81 CTGCTCCCTGCTGTGTGTTGGAGTCTGCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTTGCATGAAGAATCTGC 100

Db 141 ACAAGGCAAGCTTGACCGACAATTTGCATGAAGAATCTGC 180

## RESULT 8

ABZ23249  
ID ABZ23249 standard; DNA; 2245 BP.

AC ABZ23249;

XX 24-MAR-2003 (first entry)

DT Lac repressor operated p21-expression cassette and RSV-LTR promoter.

DE Lac repressor; p21; RSV; LTR promoter; cell cycle inhibitor protein;  
KW protein production; anchorage-independent producer cell line; ss.

XX Synthetic.

OS Key Location/Qualifiers  
FH Promoter 1..563

FT /\*tag= a

FT /note= "RSV-LTR promoter"

FT intron 564..1051

FT /\*tag= b

FT /note= "SV40 small t antigen intron"

FT misc\_feature 1052..1907

FT /\*tag= c

FT /note= "p21 coding sequence"

FT polyA\_signal 1908..2245

FT /\*tag= d

FT /note= "thymidine kinase polyA site"

XX WO200299100-A2.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-EP006054.

XX 01-JUN-2001; 2001GB-00013318.

XX (LONZ ) LONZA BIOLOGICS PLC.

XX Al-Rubeai M, Shuttleworth J;

XX WPI; 2003-148669/14.

XX Producing recombinant protein, particularly for maximizing or enhancing  
PT e.g. therapeutic protein production, by co-expressing protein with  
PT recombinant cell cycle inhibitor protein (p21) in producer cell line.

XX Example 1; Page 15-16; 33pp; English.

XX The present sequence represents a lac repressor operated p21-expression  
CC cassette comprising the Rous sarcoma virus (RSV)-LTR promoter. p21 is a  
CC cell cycle inhibitor protein. The present sequence is used to produce  
CC vectors for use in the method of the invention. The specification  
CC describes a method for producing a protein, preferably a recombinant  
CC protein, in a mammalian anchorage-independent producer cell line. The  
CC method comprises co-expressing with the protein in the producer cell line  
CC a recombinant cell cycle inhibitor protein (preferably p21). The method  
CC is useful for producing a recombinant protein in a producer cell line.  
CC This is particularly useful for maximizing or enhancing the production of  
CC e.g. therapeutic proteins at an industrial scale

XX

SQ Sequence 2245 BP; 532 A; 555 C; 625 G; 533 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 8; Length 2245;

Best Local Similarity 100.0%; Pred. No. 1.3e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTTGGAGTCTGCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 60  
Db 46 CTGCTCCCTGCTGTGTGTTGGAGTCTGCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 105

Qy 61 ACAAGGCAAGCTTGACCGACAATTTGCATGAAGAATCTGC 100

Db 106 ACAAGGCAAGCTTGACCGACAATTTGCATGAAGAATCTGC 145

## RESULT 9

ADM41036  
ID ADM41036 standard; DNA; 2294 BP.

XX ADM41036;

XX 17-JUN-2004 (first entry)

DT Cytomegalovirus nucleotide sequence SEQ ID NO:4.

DE engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX Cytomegalovirus.

OS WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschoner WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,  
PT useful in producing chimeric mammals, comprises selectively destroying  
PT native cells in a tissue of a fetal non-human mammal host.

XX Disclosure; SEQ ID NO 4; 48pp; English.

XX The present invention describes a method for engrafting foreign  
CC replacement cells within a foetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a foetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the fetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2294;

Best Local Similarity 100.0%; Pred. No. 1.3e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60  
 |||  
 Db 81 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 140  
 |||  
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100  
 |||  
 Db 141 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 180  
 |||

RESULT 10  
 AAD02037  
 ID AAD02037 standard; DNA; 2426 BP.  
 XX  
 AC AAD02037;  
 XX  
 DT 11-SEP-2003 (revised)  
 DT 26-MAR-2001 (first entry)  
 XX  
 XX Plasmid pNG3/RC/CPG2-Thy1 comprising CPG2 DNA with rat thy1 gene.  
 XX Carboxypeptidase G2; CPG2; gene directed enzyme prodrug therapy; GDEPT;  
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;  
 KW plasmid; ds.  
 XX Rattus sp.  
 OS Bacteria.  
 OS Chimeric.  
 XX WO200066752-A2.  
 FN 09-NOV-2000.  
 PD 28-APR-2000; 2000WO-GB001640.  
 XX 01-MAY-1999; 99GB-00010077.  
 PR (ASTR ) ASTRAZENECA AB.  
 PA (UTMA-) UNIV VICTORIA MANCHESTER.  
 XX Castro MG, Emery SC, Lowenstein PR;  
 PI WPI; 2001-015983/02.  
 DR  
 XX  
 XX Gene directed enzyme prodrug therapy using post translational  
 PT glycosylphosphatidylinositol addition to prodrug activating enzyme to  
 PT enable anchorage of enzyme at cell surface for cancer therapy.  
 XX  
 PS Example 1e; Page 59-60; 60pp; English.  
 XX  
 CC The present invention relates to a gene directed enzyme prodrug therapy  
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)  
 CC addition to a prodrug activating enzyme which enables anchorage of the  
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred  
 CC prodrug activating enzyme. The invention also relates to an expression  
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the  
 CC surface of a mammalian cell. The expression vector comprise  
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of  
 CC activating a prodrug, and a post-translational GPI addition motif. The  
 CC expression vector is useful in the manufacture of a medicament for cancer  
 CC therapy in a mammalian host. The present DNA sequence is a plasmid  
 CC pNG3/RC/CPG2-Thy1 comprising CPG2 nucleic acid sequence with the last  
 CC exon of rat thy 1 gene at its 3' end. (Updated on 11-SEP-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 2426 BP; 557 A; 705 C; 668 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2426;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60  
 |||  
 Db 69 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 128  
 |||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100  
 |||  
 Db 129 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 168  
 |||  
 RESULT 11  
 AAD02036  
 ID AAD02036 standard; DNA; 2427 BP.  
 XX  
 AC AAD02036;  
 XX  
 DT 11-SEP-2003 (revised)  
 DT 26-MAR-2001 (first entry)  
 XX  
 XX Plasmid pNG3/RC/CPG2(Q3)-Thy1 comprising CPG2 variant with rat thy1 gene.  
 XX Carboxypeptidase G2; CPG2; gene directed enzyme prodrug therapy; GDEPT;  
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;  
 KW CPG2(Q3) variant; plasmid; ds.  
 XX Rattus sp.  
 OS Bacteria.  
 OS Chimeric.  
 XX WO200066752-A2.  
 FN 09-NOV-2000.  
 PD 28-APR-2000; 2000WO-GB001640.  
 XX 01-MAY-1999; 99GB-00010077.  
 PR (ASTR ) ASTRAZENECA AB.  
 PA (UTMA-) UNIV VICTORIA MANCHESTER.  
 XX Castro MG, Emery SC, Lowenstein PR;  
 PI WPI; 2001-015983/02.  
 DR  
 XX  
 XX Gene directed enzyme prodrug therapy using post translational  
 PT glycosylphosphatidylinositol addition to prodrug activating enzyme to  
 PT enable anchorage of enzyme at cell surface for cancer therapy.  
 XX  
 PS Example 1e; Page 59; 60pp; English.  
 XX  
 CC The present invention relates to a gene directed enzyme prodrug therapy  
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)  
 CC addition to a prodrug activating enzyme which enables anchorage of the  
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred  
 CC prodrug activating enzyme. The invention also relates to an expression  
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the  
 CC surface of a mammalian cell. The expression vector comprise  
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of  
 CC activating a prodrug, and a post-translational GPI addition motif. The  
 CC expression vector is useful in the manufacture of a medicament for cancer  
 CC therapy in a mammalian host. The present DNA sequence is a plasmid  
 CC pNG3/RC/CPG2(Q3) comprising CPG2 variant CPG2(Q3) and the last exon of  
 CC rat Thy-1 at the 3' end. (Updated on 11-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 2427 BP; 555 A; 706 C; 670 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2427;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60  
 |||  
 Db 70 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 129  
 |||  
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100  
 |||  
 Db 130 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 169  
 |||



```
RESULT 12
AAT62937
ID AAT62937 standard; DNA; 3400 BP.
XX
AC AAT62937;
XX
DT 17-OCT-2003 (revised)
DT 16-JUN-1997 (first entry)
XX
DE 3F4 human G2/G4 chimeric antibody expression plasmid insert.
XX
KW Xenotransplantation; graft rejection; cell interaction; pig;
KW vascular cell adhesion molecule; VCAM; monoclonal antibody;
KW chimeric antibody; diagnosis; ss.
XX
OS Homo; sapiens.
OS Mus sp.
OS Chimeric.
XX
FH Key
FT exon Location/Qualifiers
    903..1055
    /*tag= a
FT intron
    1056..1285
    /*tag= b
FT exon
    1286..2055
    /*tag= c
FT intron
    /codon_start= 1350
    2056..2447
    /*tag= d
FT exon
    2448..2483
    /*tag= e
FT intron
    2484..2601
    /*tag= f
FT exon
    2602..2928
    /*tag= g
FT intron
    2929..3025
    /*tag= h
FT exon
    3026..3348
    /*tag= i
XX
PN WO9711971-A1.
XX
PD 03-APR-1997.
XX
PF 27-SEP-1996; 96WO-US015575.
XX
PR 28-SEP-1995; 95US-0004489P.
PR 26-SEP-1996; 96US-00004489.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;
XX
DR WPI; 1997-212855/19.
DR P-PSDB; AAW14940.
XX
XX Antibodies binding to porcine but not human cell interaction proteins -
PT useful to treat and assay for rejection of xenografted porcine organs,
PT tissues or cells.
XX
PS Disclosure; Page 58-61; 105pp; English.
XX
CC A DNA sequence (AAT62937) comprises a 3F4 human G2/G4 (see also AAT62936)
CC chimeric antibody expression plasmid insert sequence. The chimeric
CC antibody (AAW14940) is specific for porcine vascular cell adhesion
CC molecule (VCAM) and is useful for diagnosing human rejection of porcine
CC xenotransplants and for improving xenotransplantation of porcine cells,
CC tissues and organs into human recipients. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 3400 BP; 759 A; 1012 C; 909 G; 720 T; 0 U; 0 Other;
```

```
Query Match . 100.0%; Score 100; DB 2; Length 3400;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGGTGCTGCTAGTGTGCGGAGCAGCAAAATTTAAGCTACA 60
Db 148 CTGCTCCCTGCTTGTGTTGGAGGTGCTGCTAGTGTGCGGAGCAGCAAAATTTAAGCTACA 207
Qy 61 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 100
Db 208 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 247

RESULT 13
AAT62932
ID AAT62932 standard; DNA; 3400 BP.
XX
AC AAT62932;
XX
DT 17-OCT-2003 (revised)
DT 16-JUN-1997 (first entry)
XX
DE 2A2 human G2/G4 chimeric antibody expression plasmid insert.
XX
KW Xenotransplantation; graft rejection; cell interaction; pig;
KW vascular cell adhesion molecule; VCAM; monoclonal antibody;
KW chimeric antibody; diagnosis; ss.
XX
OS Homo; sapiens.
OS Mus sp.
OS Chimeric.
XX
FH Key
FT exon Location/Qualifiers
    903..1055
    /*tag= a
FT intron
    1056..1285
    /*tag= b
FT exon
    1286..2020
    /*tag= c
FT intron
    /codon_start= 1318
    2021..2412
    /*tag= d
FT exon
    2413..2448
    /*tag= e
FT intron
    2449..2566
    /*tag= f
FT exon
    2567..2983
    /*tag= g
FT intron
    2984..2990
    /*tag= h
FT exon
    2991..3313
    /*tag= i
XX
PN WO9711971-A1.
XX
PD 03-APR-1997.
XX
PF 27-SEP-1996; 96WO-US015575.
XX
PR 28-SEP-1995; 95US-0004489P.
PR 26-SEP-1996; 96US-00004489.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;
XX
DR WPI; 1997-212855/19.
DR P-PSDB; AAW14934.
XX
XX Antibodies binding to porcine but not human cell interaction proteins -
PT useful to treat and assay for rejection of xenografted porcine organs,
PT tissues or cells.
```

```

XX PS Disclosure; Page 44-47; 105pp; English.
XX CC
XX CC A DNA sequence (AAT62932) comprises a 2A2 human G2/G4 (see also AAT62931)
XX CC chimeric antibody expression plasmid insert sequence. The chimeric
XX CC antibody (AAW14934) is specific for porcine vascular cell adhesion
XX CC molecule (VCAM) and is useful for diagnosing human rejection of porcine
XX CC xenotransplants and for improving xenotransplantation of porcine cells,
XX CC tissues and organs into human recipients. (Updated on 17-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 3400 BP; 776 A; 993 C; 899 G; 732 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 3400;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGGACGACAAATTTAAGCTACA 60
Db 148 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGGACGACAAATTTAAGCTACA 207

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCAATGCAAGAAATCTGC 100
Db 208 ACAAGGCAAGCTTGACCGACAATTGCAATGCAAGAAATCTGC 247

RESULT 14
AAV40006
ID AAV40006 standard; DNA; 3853 BP.
AC AAV40006;
XX
XX 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
XX
XX Plasmid pCTM.
XX E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
XX Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH promoter 209..864
FT /tag= a
FT /note= "CMV promoter"
FT misc_feature 907..1131
FT /tag= b
FT /function= "tripartite leader sequence"
FT promoter 1132..1149
FT /tag= c
FT /note= "SP6 promoter"
FT misc_feature 1679..3853
FT /tag= d
FT /note= "pUC19 backbone H3 to AatII"
FT CDS complement(2857..3717)
FT /tag= e
FT /note= "AMP-ORF"
XX
XX WO9821228-A1.
FN
XX 22-MAY-1998.
PD
XX 13-NOV-1997; 97WO-US021821.
PF

```

---

```

XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX PA (CANJ-) CANJI INC.
XX PI Antelman D, Gregory RJ, Wills KN;
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 4; 91pp; English.
XX
XX This is the nucleotide sequence of pCTM, a plasmid which contains a CMV
XX promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters,
XX and a multiple cloning site with a bovine growth hormone polyA site and
XX downstream SV40 polyA site. It has been used as a vector for the
XX expression of fusion proteins of the invention that comprise
XX retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX (see AAW62464). Such fusion proteins, particularly expressed from gene
XX therapy vectors, are used to treat hyperproliferative conditions,
XX specifically cancer (particularly of the bladder) or restenosis. They are
XX more effective in repressing transcription of the E2F promoter than RB
XX alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 3853;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGGACGACAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGGACGACAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCAATGCAAGAAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCAATGCAAGAAATCTGC 180

RESULT 15
AAT90695
ID AAT90695 standard; DNA; 3925 BP.
XX
XX AC AAT90695;
XX
XX 05-JAN-1998 (first entry)
DT
XX
XX Plasmid CMV10A1 coding sequence.
XX
XX Packaging-deficient construct; viral gag-pol gene; packaging cell line;
XX moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
XX gene therapy; human cytomegalovirus; promoter; ss.
XX
XX Synthetic.
OS
XX WO9708330-A1.
XX
XX 06-MAR-1997.
PD
XX
XX 23-AUG-1996; 96WO-GB002061.
PF
XX
XX 23-AUG-1995; 95GB-00017263.
PR
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA
XX Collins MKL, Weiss RA, Takeuchi Y, Cosset F;
XX WPI; 1997-179287/16.
XX
XX

```

PT Selectable retroviral packaging cell lines and expression constructs -  
PT comprise selectable gene downstream of gene of interest, are selectable  
PT due to the in-efficiency associated with translation re-initiation.  
XX  
PS  
PS Claim 23; Fig 13; 79pp; English.  
XX  
CC This sequence represents the recombinant expression plasmid CMV10A. This  
CC sequence is a packaging-deficient construct having a viral env gene (in  
CC this case from moloney murine leukaemia virus under hCMV promoter  
CC control) and a selectable marker (SM). It is an example of a recombinant  
CC expression vector (REV) of the invention, used to create a packaging cell  
CC line. The REV's of the invention comprise a gene of interest (GOI) and a  
CC SM gene. The SM gene is arranged downstream of the GOI and a GOI  
CC associated stop codon is spaced from a start codon of the SM gene to  
CC ensure that the SM protein is expressed as a result of translation  
CC reinitiation. The cell lines are transformed with two REV's, both are  
CC replication deficient, one contains the viral gag-pol gene, the other the  
CC viral env gene. By using helper constructs, such as the REV's, which are  
CC directly selectable and which provide for high expression of the viral  
CC gene, high titre retroviral vectors may be obtained. The packaging cell  
CC lines are useful for gene therapy. Prior packaging cell lines using full  
CC length retroviral genomes as helper genomes were isolated by  
CC cotransfecting them with plasmids encoding selectable markers. However,  
CC the helper functions can be lost during the passages of the cells in  
CC culture and the current packaging systems provide limited titres of  
CC infectious retroviral vectors. Co-transfection with a plasmid encoding a  
CC SM does not directly select the best gag-pol-env-expressing cells. The  
CC new retroviral packaging cell lines overcome these problems  
XX  
SQ Sequence 3925 BP; 963 A; 1001 C; 959 G; 998 T; 0 U; 4 Other;

Query Match 100.0%; Score 100; DB 2; Length 3925;  
Best Local Similarity 100.0%; Pred. No. 1.6e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAGCTACA 60  
Db 70 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAGCTACA 129  
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAAATCTGC 100  
Db 130 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAAATCTGC 169

Search completed: July 14, 2005, 07:01:45  
Job time : 143.038 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_1\_100

Perfect score: 100

Sequence: 1 ctgtccctgtgtgtgtt.....caattgcatgaagaattgc 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	602	8	B67169 CPG0047A Cp
2	30.4	30.4	829	4	B1333630 602997459
3	30.2	30.2	823	6	CD655614 AGENCOURT
4	29.8	29.8	1165	8	CC242469 CH261-11F
5	29.4	29.4	401	6	CB387202 OSTF076E6
6	29.4	29.4	531	5	BQ310441 MR0-BT450
7	29.4	29.4	754	8	AQ946479 Sheared D
8	29.2	29.2	426	9	CC888514 SALK 1519
9	29	29.0	657	7	CK086063 RG11-C07
10	29	29.0	877	9	AK225351 Tetraodon
11	28.6	28.6	340	7	F32722 HSPD25699 H
12	28.6	28.6	408	1	AA962465 oc9le05.s
13	28.6	28.6	436	7	CN386744 328755673
14	28.6	28.6	514	6	CB161201 K-EST0221
15	28.6	28.6	530	6	CB161182 K-EST0220
16	28.6	28.6	534	2	AW500392 UI-HF-BN0
17	28.6	28.6	550	1	AA984313 am83h04.s
18	28.6	28.6	555	7	CR537056 DKFZ0459D
19	28.6	28.6	583	7	CN386728 170005326
20	28.6	28.6	585	4	BG993413 MR3-HT099
21	28.6	28.6	625	4	B1113747 602860946
22	28.6	28.6	626	7	CN386724 170006000
23	28.6	28.6	641	2	AW955076 EST367146
24	28.6	28.6	651	7	CN386730 170005316

C 25 28.6 28.6 680 7 CN386731 170005313  
 C 26 28.6 28.6 682 7 CN386746 170005999  
 C 27 28.6 28.6 687 7 CN386717 170004551  
 C 28 28.6 28.6 710 2 BF309673 601891808  
 C 29 28.6 28.6 724 7 CN386780 170005318  
 C 30 28.6 28.6 770 6 CD654097 AGENCOURT  
 C 31 28.6 28.6 777 7 CN386742 170005314  
 C 32 28.6 28.6 801 6 CD656906 AGENCOURT  
 C 33 28.6 28.6 814 4 BG820298 602782110  
 C 34 28.6 28.6 823 6 CB996419 AGENCOURT  
 C 35 28.6 28.6 827 4 BG387788 602412672  
 C 36 28.6 28.6 830 6 CD643822 AGENCOURT  
 C 37 28.6 28.6 842 5 BUI70446 AGENCOURT  
 C 38 28.6 28.6 851 6 CB993607 AGENCOURT  
 C 39 28.6 28.6 852 5 BQ222104 AGENCOURT  
 C 40 28.6 28.6 856 6 CD656102 AGENCOURT  
 C 41 28.6 28.6 909 5 BUI89860 AGENCOURT  
 C 42 28.6 28.6 1025 4 BM477450 AGENCOURT  
 C 43 28.4 28.4 579 6 CB239722 RSH15G08  
 C 44 28.4 28.4 733 7 CF667137 RTCTN1.28  
 C 45 28.4 28.4 1015 6 BY703355 BY703355

## ALIGNMENTS

RESULT 1  
 LOCUS B67169 602 bp DNA linear GSS 12-MAY-2000  
 DEFINITION CPG0047A CpioWagDNA2 Cryptosporidium parvum genomic, genomic survey sequence.  
 ACCESSION B67169  
 VERSION B67169.1 GI:2642750  
 KEYWORDS GSS.  
 SOURCE Cryptosporidium parvum  
 ORGANISM Cryptosporidium parvum  
 REFERENCE 1 (bases 1 to 602)  
 AUTHORS Strong, W.B. and Nelson, R.G.  
 TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
 JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
 MEDLINE 20183851  
 PUBMED 10717299  
 COMMENT Contact: Nelson, R. G.  
 Depts. of Medicine & Pharmaceutical Chemistry  
 San Francisco General Hospital-University of California, San Francisco  
 Box 0811, San Francisco, CA 94143-0811, USA  
 Tel: 415 206 8846  
 Fax: 415 206 3353  
 Email: malariad@itsa.ucsf.edu  
 Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.  
 Seq primer: T7  
 Class: Shotgun  
 High quality sequence stop: 602.  
 Location/Qualifiers  
 1..602  
 /organism="Cryptosporidium parvum"  
 /mol\_type="genomic DNA"  
 /strain="IOWA"  
 /db\_xref="taxon:5807"  
 /lab\_host="E. coli XL2 Blue MRF"  
 /clone\_lib="CpioWagDNA2"  
 /note="vector: PCR-script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

#### ORIGIN

Query Match 100.0%; Score 100; DB 8; Length 602;  
 Best Local Similarity 100.0%; Pred. No. 8e-24;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
 |||||  
 Db 41 CTGCTCCCTGCTTGTGTGGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 100  
 |||||

Qy 61 ACAAGGCAAGCTTGACGACAATTGATGAAGAAATCTGC 100  
 |||||  
 Db 101 ACAAGGCAAGCTTGACGACAATTGATGAAGAAATCTGC 140  
 |||||

#### RESULT 2

BI333630/c  
 LOCUS  
 DEFINITION 602997459P1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5139470 5',  
 mRNA sequence.

ACCESSION BI333630  
 VERSION BI333630.1 GI:15018287  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 829)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM11343 row: f column: 15  
 High quality sequence stop: 788.  
 Location/Qualifiers

#### FEATURES

source  
 1..829  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5139470"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Site 3: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

#### ORIGIN

Query Match 30.4%; Score 30.4; DB 4; Length 829;  
 Best Local Similarity 67.2%; Pred. NO. 15;  
 Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 19 TTGGAGGTCGCTAGTAGTCGCGAGCAAAATTTAAGCTACAACAGCAAGGCTTGACC 78  
 |||||

Db 217 TTGGCGCTCCCAAGATTGTTGGTGAGCACAATTCAAGTGTGCTGCTGGGAGTCTGACT 158  
 Qy 79 GACA 82  
 |||||  
 Db 157 GACA 154

RESULT 3  
 CD655614/c  
 LOCUS  
 DEFINITION CD655614 823 bp mRNA linear EST 18-JUN-2003  
 (Long) Homo sapiens cDNA clone IMAGE:30424285 5', mRNA sequence.

ACCESSION CD655614  
 VERSION CD655614.1 GI:31896113  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 823)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
 cDNA Library Preparation: Yulan Piao and Minoru Ko  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC c lone distribution information  
 can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM506 row: k column: 14  
 High quality sequence stop: 640.  
 Location/Qualifiers

#### FEATURES

source  
 1..823  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30424285"  
 /tissue\_type="Embryonic Stem cells"  
 /cell\_line="WA01"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"  
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Site 3: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLIT3, SSEA-1, TUBB3, NES, GFAP, and ROMEs. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGTCGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase



```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-BT4502-
220601-202-b02&t3=2001-06-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 48
High quality sequence stop: 530.
FEATURES
Location/Qualifiers
source 1..531
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT4502"
/notes="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORSITES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 29.4%; Score 29.4; DB 5; Length 531;
Best Local Similarity 60.8%; Pred. No. 31;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 22 GAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACACAGCGAGGCTTGACCGAC 81
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 152 GAGCTAACTGAATGTATGGGAGCAGCATTTAAACATATCTTAGTCAGGACGAGATGGG 93
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 82 AATTCATGAAGAATCTGC 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 AAGTAAGTGAAGATAGGC 74
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
AQ946479/c 754 bp DNA linear GSS 27-JAN-2000
LOCUS Sheared DNA-4906.TR Sheared DNA Trypanosoma brucei genomic clone
DEFINITION Sheared DNA-4906, genomic survey sequence.
ACCESSION AQ946479
VERSION AQ946479.1 GI:6769744
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 754)
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,
Doneison,J., Fraser,C. and Adams,M.
TITLE Determination of clone end sequences from Trypanosoma brucei GUTat
10.1 sheared DNA library
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: Sheared DNA-4906.TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tdb/mdb/tbdb/.
Seq primer: M13-Reverse
Class: Shotgun.
FEATURES
Location/Qualifiers
source 1..754
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927/4 GUTat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-4906"
/clone_lib="Sheared DNA"
/notes="Vector: pUC18; Site 1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In genome sequencing: A Practical
Approach, eds. M. Vaudin and B. Borell, Oxford University
Press, 1999)."
ORIGIN
Query Match 29.4%; Score 29.4; DB 8; Length 754;
Best Local Similarity 70.9%; Pred. No. 33;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 45 CAAATTTAAGCTACACAGCAAGGCTTGACCGCAATTCATGAAGATCTG 99
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 165 CACAAATTAAGCTTCAAAAAGGCAAGCCTGCACAGCATGTGCAGAAAGAAAGTG 111
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
CC888514/c 426 bp DNA linear GSS 31-JUL-2003
LOCUS SALK_151964.54.50.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_151964.54.50.x, genomic
survey sequence.
ACCESSION CC888514
VERSION CC888514.1 GI:33365229
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 426)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker

```



Salk Institute Genomic Analysis Laboratory (SIGnal)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: eckergaalk.edu  
This is single pass sequence recovered from the left border of  
TDNA. This sequence lies within an annotated intron of AL1903960.  
Class: TDNA tagged.

## FEATURES

source

Location/Qualifiers  
1..426  
/organism="Arabidopsis thaliana"  
/mol\_type="Genomic DNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="SALK151964.54.50.x"  
/note="PCR was performed on Arabidopsis thaliana lines  
each of which contains one or more TDNA insertion  
elements. The resultant fragment for each line was  
directly sequenced to determine the genomic sequence at  
the site of insertion. Details of the protocols used can  
be found at [http://signal.salk.edu/cdna\\_protocols.html](http://signal.salk.edu/cdna_protocols.html)"

## ORIGIN

Query Match 29.2%; Score 29.2; DB 9; Length 426;  
Best Local Similarity 64.2%; Pred. No. 35;  
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 30 TGAGTAGTGGCGGACCAAAATTTAGCTACACAGGCAAGCTTGACGCAATTGCAAT 89  
DB 76 TTATTAGTTGGTGTTCAGANTTTAGCATCATCAAGACTTGACCTACAGAAACAT 17

QY 90 GAAGAAAT 96

DB 16 GAAAAAT 10

## RESULT 9

CK086063/C  
LOCUS RG11\_C07 Cucumber leaf Cucumis sativus cDNA, mRNA linear EST 01-DEC-2003

ACCESSION CK086063  
VERSION CK086063.1 GI:38571123

KEYWORDS EST.

SOURCE Cucumis sativus (cucumber)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

## REFERENCE

AUTHORS Grunet,R. and McGrath,M.

TITLE Development of genomic tools for cucumber (Cucumis sativus L.)

JOURNAL Unpublished (2003)

COMMENT Contact: Rebecca Grunet

Rebecca Grunet

Michigan State University

Horticulture Department, Michigan State University, East Lansing,

MI 48824, USA

Tel: 517 353 0890

Fax: 517 355 5191 x431

Email: grunet@msu.edu

Plate: RG11 row: C column: 07.

Location/Qualifiers

1..657

/organism="Cucumis sativus"

/mol\_type="mRNA"

/strain="Straight 8"

/db\_xref="taxon:3659"

/sex="monoecious"

/clone\_lib="Cucumber leaf"

/note="Vector: pAD-GAL4; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 29.0%; Score 29; DB 7; Length 657;  
Best Local Similarity 63.8%; Pred. No. 44;  
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 31 GAGTAGTGGCGGACCAAAATTTAGCTACACAGGCAAGCTTGACGCAATTGCAATG 90

DB 267 GAGAGTGCAAGAACCAAACTGAAGCCAGATGAAGAGAGGCTTGCAAGGTCAAG 208

QY 91 AAGAATCTG 99

DB 207 AAAAAATTTG 199

## RESULT 10

CNS032VI/C

LOCUS CNS032VI/C

DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone

207F02 of library G from Tetraodon nigroviridis, genomic survey

sequence.

ACCESSION AL225351.1 GI:7884242

VERSION AL225351.1

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1

AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,

Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,

Saurin,W. and Weissenbach,J.

Estimate of human gene number provided by genome-wide analysis

using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

20296633

REFERENCE 2

AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,

Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,

Saurin,W., Bernot,A. and Weissenbach,J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)

20359837

PUBMED 10899143

REFERENCE 3 (bases 1 to 877)

Genoscope.

Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))

- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

1..877

/organism="Tetraodon nigroviridis"

/mol\_type="genomic DNA"

/db\_xref="taxon:99883"

/clone="207F02"

/clone\_lib="G"

/note="Genoscope sequence ID : COAG207DC01SP1-end :

PUC-Ori"

ORIGIN

Query Match 29.0%; Score 29; DB 9; Length 877;

Best Local Similarity 63.8%; Pred. No. 47;

Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 14 GTGTGTTGGAGTGGCTGAGTAGTGGCGGACGACAAAATTTAGCTACACAGGCAAGGCT 73

```

Db      199  GGGTCTGGGGCGCTGCTCGGTGCTGAAGCATAAACATCTACACATGGAAAGTTA 140
Qy      74  TGACCGACA 82
Db      139  TGAAGAGAGA 131

RESULT 11
F32722/c
LOCUS      F32722
DEFINITION HSPD25699 HM3 Homo sapiens cDNA clone s3000037G06, mRNA sequence.
ACCESSION F32722
VERSION    F32722.1 GI:4818348
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 340)
            Landfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,
            Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
            Identification of 4370 expressed sequence tags from a
            3'-end-specific cDNA library of human skeletal muscle by DNA
            sequencing and filter hybridization
            Genome Res. 6 (1), 35-42 (1996)
JOURNAL    96276048
MEDLINE    8691137
COMMENT    Contact: Valle G.
            CRIBI Biotechnology Centre
            University of Padua
            Via Trieste 75, 35121 Padua, Italy
            ABI Chromatograms and other information are available on WWW at
            http://group.bio.unipd.it.
            Location/Qualifiers
FEATURES   source
            1..340
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="s3000037G06"
            /sex="female"
            /tissue_type="pectoral muscle (after mastectomy)"
            /clone_lib="HM3"
            /note="Vector: pcDNAII (Invitrogen); Site 1: BstXI;
            Site 2: NotI; The library is not constructed by G.
            Lanfranchi. This library is not subtracted nor normalized.
            The first strand cDNA was primed with a biotinylated
            oligo-dT-NotI primer
            (5'-biotin-AACCGCTCGAGCGCGCTTTT-3'). The
            ds cDNA was sonicated and size-selected in the range
            350-550 bp. The 3' specific fragments were selected by
            streptavidin coated magnetic beads, ligated to
            non-palindromic BstXI adapters, NotI digested and
            directionally cloned into BstXI-NotI cut pcDNAII vector."

ORIGIN
Query Match      28.6%; Score 28.6; DB 7; Length 340;
Best Local Similarity 61.3%; Pred. No. 55;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy      22  GAGTCGCTGAGTAGTCGCGAGCAAAATTTAAGCTACAACAGCGGCTTGACCGAC 81
Db      125  GAGCTAACTGAATGTGTATGGAGCAGCATTTAACAATATTCCTAGTCAAGGACAGGATGGG 66
Qy      82  AATTGCATGAAGAAT 96
Db      65  AAGTAAGTGAAGAAT 51

RESULT 12
AA962465
LOCUS      AA962465
DEFINITION o091e05.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1573568 3',

```

```

mRNA sequence.
AA962465
VERSION    AA962465.1 GI:3134629
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 408)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LINL at:
            www-bio.lnl.gov/bbrp/image/image.html
            Insert Length: 1045 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
            Location/Qualifiers
FEATURES   source
            1..408
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1573568"
            /tissue_type="2 pooled tumors (clear cell type)"
            /lab_host="DHI08"
            /clone_lib="NCI CGAP_Kid5"
            /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
            a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAAATCGCGCGCAATATTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match      28.6%; Score 28.6; DB 1; Length 408;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy      22  GAGTCGCTGAGTAGTCGCGAGCAAAATTTAAGCTACAACAGCGGCTTGACCGAC 81
Db      243  GAGCTAACTGAATGTGTATGGAGCAGCATTTAACAATATTCCTAGTCAAGGACAGGATGGG 302
Qy      82  AATTGCATGAAGAAT 96
Db      303  AAGTAAGTGAAGAAT 317

RESULT 13
CN386744/c
LOCUS      CN386744
DEFINITION 328755673 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN386744
VERSION     CN386744.1 GI:47374339
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens

REFERENCE
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 436)
            Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
            Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
            Lebkowski, J and Stanton, L.W.

```

TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

COMMENT Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 436 Std Error: 0.00.

FEATURES  
source Location/Qualifiers  
1..436  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN  
Query Match 28.6%; Score 28.6; DB 7; Length 436;  
Best Local Similarity 61.3%; Pred.No.57;  
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
QY 22 GAGTGCCTGAGTAGTCGGCAGCAAAATTTAAGCTACACAGGCAAGCGCTTGACCGAC 81  
Db 133 GAGCTAAGTGAATGTGTGGGAGCATTTAAATCTCCTAGTCAAGGACAGGATGGG 74  
QY 82 AATTGCATGAAGAAT 96  
Db 73 AAGTAAGTGAAGAAT 59

RESULT 14  
LOCUS CB161201  
DEFINITION K-EST0221011 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-H03  
5', mRNA sequence.  
ACCESSION CB161201  
VERSION CB161201.1 GI:28147327  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 514)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsaung@mail.kribb.re.kr  
Plate: 16 row: H column: 03  
High quality sequence stop: 514.  
Location/Qualifiers  
1..514  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L18POOL1n1-16-H03"  
/cell\_line="SNU-354+Cho-CK+Choi-CK+HLK-3"  
/lab\_host="Tcpl0f"

Qy 82 AATTGCATGAGAAT 96  
|||  
Db 321 AAGTAAGTGAGAAT 335  
|||

Search completed: July 14, 2005, 23:23:14  
Job time : 960.146 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_3565\_3665  
Perfect score: 101  
Sequence: 1 gcgtgacgctacacttgc.....ttccccgcgaagctctaaat 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	299	6	CQ815667
C 2	101	100.0	460	6	I08196
C 3	101	100.0	461	6	I05488
C 4	101	100.0	461	6	I08820
C 5	101	100.0	461	6	I16794
C 6	101	100.0	470	6	A60961
C 7	101	100.0	470	6	A60977
C 8	101	100.0	470	6	AR369178
C 9	101	100.0	470	6	AR369188
C 10	101	100.0	470	6	AR476628
C 11	101	100.0	470	6	AR476638
C 12	101	100.0	470	6	AR487254
C 13	101	100.0	470	6	AR487264
C 14	101	100.0	472	6	AR160383
C 15	101	100.0	472	6	BD194797
C 16	101	100.0	472	6	AX482611
C 17	101	100.0	698	6	A85395
C 18	101	100.0	698	6	AR154888
C 19	101	100.0	698	6	E65413

C 20	101	100.0	729	6	BD237245	BD237245 Metastati
C 21	101	100.0	729	6	AR240810	AR240810 Sequence
C 22	101	100.0	730	6	BD237246	BD237246 Metastati
C 23	101	100.0	730	6	AR240811	AR240811 Sequence
C 24	101	100.0	757	6	BD221144	BD221144 Human gen
C 25	101	100.0	801	6	AR370661	AR370661 Sequence
C 26	101	100.0	825	6	AX284008	AX284008 Sequence
C 27	101	100.0	858	6	AX752973	AX752973 Sequence
C 28	101	100.0	902	6	AR505921	AR505921 Sequence
C 29	101	100.0	910	9	LDO244004	AJ244004 Lepilemur
C 30	101	100.0	1217	6	AR036903	AR036903 Sequence
C 31	101	100.0	1217	6	AR141142	AR141142 Sequence
C 32	101	100.0	1217	6	AR181917	AR181917 Sequence
C 33	101	100.0	1218	5	AJ719748	AJ719748 Gallus ga
C 34	101	100.0	1695	12	AB003139	AB003139 Arabidops
C 35	101	100.0	2000	6	I01987	I01987 Sequence 7
C 36	101	100.0	2029	12	AY733067	AY733067 Cloning v
C 37	101	100.0	2175	12	AY733070	AY733070 Cloning v
C 38	101	100.0	2322	6	AX670965	AX670965 Sequence
C 39	101	100.0	2326	12	AF143508	AF143508 Cloning v
C 40	101	100.0	2380	6	A60985	A60985 Sequence 29
C 41	101	100.0	2380	6	AR369195	AR369195 Sequence
C 42	101	100.0	2380	6	AR476645	AR476645 Sequence
C 43	101	100.0	2380	6	AR487271	AR487271 Sequence
C 44	101	100.0	2383	12	CVU50331	U50331 Cloning vec
C 45	101	100.0	2402	6	AX670963	AX670963 Sequence

#### ALIGNMENTS

RESULT 1  
CQ815667/c  
LOCUS CQ815667 299 bp DNA linear PAT 03-JUN-2004  
DEFINITION Sequence 65 from Patent WO2004042036.  
ACCESSION CQ815667  
VERSION CQ815667.1 GI:48144221

KEYWORDS  
SOURCE  
ORGANISM  
Zygosaccharomyces bailii  
Zygosaccharomyces bailii  
Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Zygosaccharomycetes;

#### REFERENCE

AUTHORS Porro,D., Branduardi,P., Valli,M. and Alberghina,L.  
TITLE Process for expression and secretion of proteins by the non-conventional yeast zygosaccharomyces bailii  
JOURNAL Patent: WO 2004042036-A 65 21-MAY-2004;

#### FEATURES

source  
Porro, Danilo (IT)  
Location/Qualifiers  
1..299  
/organism="Zygosaccharomyces bailii"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4954"

#### ORIGIN

Query Match: 100.0%; Score 101; DB 6; Length 299;  
Best Local Similarity 100.0%; Pred. No. 6.6e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCCTGACCGCTACACTTCCAGCGCCCTAGCGCCGCTTCCCTTCCTTCCTTCCT 60  
Db 156 GCCTGACCGCTACACTTCCAGCGCCCTAGCGCCGCTTCCCTTCCTTCCTTCCT 97  
Qy 61 TTCTGCGCAGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101  
Db 96 TTCTGCGCAGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 56

#### RESULT 2

LOCUS I08196 460 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0356130.  
ACCESSION I08196

```
VERSION I08196.1 GI:589091
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 460)
AUTHORS Weber, S.C., Holzschu, D.Lc. and Lalik, P.Hc.
TITLE A mobile fl phage single-strand DNA origin of replication
JOURNAL Patent: EP 0356130-A2 1 28-FEB-1990;
FEATURES Location/Qualifiers
1..460
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 460;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 60
Db 406 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 347
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 346 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 306
RESULT 3
I05488
LOCUS I05488 461 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0286200.
ACCESSION I05488
VERSION I05488.1 GI:590716
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 461)
AUTHORS Sorge, J.A.M., Huse, W.M. and Short, J.M.
TITLE DNA cloning vector with in vivo excisable plasmids
JOURNAL Patent: EP 0286200-A2 1 12-OCT-1988;
FEATURES Location/Qualifiers
1..461
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 154
RESULT 4
I08820
LOCUS I08820 461 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8805085.
ACCESSION I08820
VERSION I08820.1 GI:588470
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 461)
AUTHORS Huse, W., Sorge, J.A. and Short, J.M.
JOURNAL Patent: WO 8805085-A 1 14-JUL-1988;
FEATURES Location/Qualifiers
1..461
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 154
RESULT 5
I16794
LOCUS I16794 461 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 2 from patent US 5478731.
ACCESSION I16794
VERSION I16794.1 GI:1251702
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 461)
AUTHORS Short, J.M.
TITLE Polycos vectors
JOURNAL Patent: US 5478731-A 2 26-DEC-1995;
FEATURES Location/Qualifiers
1..461
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 154
RESULT 6
A60961
LOCUS A60961 470 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 270 from Patent WO9708320.
ACCESSION A60961
VERSION A60961.1 GI:3715496
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Knapik, A., Pack, P., Ilag, V., Ge, L., Moroney, S. and Plueckthun, A.
TITLE PROTEIN/(POLY)PEPTIDE LIBRARIES
JOURNAL Patent: WO 9708320-A 270 06-MAR-1997;
FEATURES Location/Qualifiers
1..470
/mol_type="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
```

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 60  
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 158

## RESULT 7

LOCUS A60977 470 bp DNA linear PAT 06-MAR-1998  
DEFINITION Sequence 286 from Patent WO9708320.

ACCESSION A60977  
VERSION A60977.1 GI:3715509

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified

REFERENCE 1  
AUTHORS Knappik, A., Pack, P., Ilag, V., Ge, L., Moroney, S. and Plueckthun, A.  
TITLE PROTEIN/(POLY)PEPTIDE LIBRARIES  
JOURNAL Patent: WO 9708320-A 286 06-MAR-1997;  
MORPHOSYS PROTEINOPTIMIERUNG (DE)

FEATURES  
source  
1..470  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 60  
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 158

## RESULT 8

LOCUS AR369178 470 bp DNA linear PAT 12-SEP-2003  
DEFINITION Sequence 270 from patent US 6300064.

ACCESSION AR369178  
VERSION AR369178.1 GI:34605134

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)  
AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.  
TITLE Protein/(poly)peptide libraries  
JOURNAL Patent: US 6300064-A 270 09-OCT-2001;  
FEATURES Location/Qualifiers

source  
1..470  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 60  
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 158

## RESULT 9

LOCUS AR369188 470 bp DNA linear PAT 12-SEP-2003  
DEFINITION Sequence 286 from patent US 6300064.

ACCESSION AR369188  
VERSION AR369188.1 GI:34605144

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)  
AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.  
TITLE Protein/(poly)peptide libraries  
JOURNAL Patent: US 6300064-A 286 09-OCT-2001;  
FEATURES Location/Qualifiers

source  
1..470  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 60  
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 158

## RESULT 10

LOCUS AR476628 470 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 270 from patent US 6696248.

ACCESSION AR476628  
VERSION AR476628.1 GI:47233721

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)  
AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.  
TITLE Protein/(poly)peptide libraries  
JOURNAL Patent: US 6696248-A 270 24-FEB-2004;  
FEATURES Location/Qualifiers

source  
1..470  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 60  
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 158

Db 118 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 158

RESULT 11

AR476638 LOCUS 470 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 286 from patent US 6696248.

ACCESSION AR476638

VERSION AR476638.1 GI:47233731

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Pluckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL Patent: US 6696248-A 286 24-FEB-2004;

FEATURES Location/Qualifiers

source 1..470

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;

Best Local Similarity 100.0%; Pred. No. 6.1e-16;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 60

Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 117

Qy 61 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 101

Db 118 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 158

RESULT 12

AR487254 LOCUS 470 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 270 from patent US 6706484.

ACCESSION AR487254

VERSION AR487254.1 GI:47252205

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Pluckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL Patent: US 6706484-A 270 16-MAR-2004;

FEATURES Location/Qualifiers

source 1..470

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;

Best Local Similarity 100.0%; Pred. No. 6.1e-16;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 60

Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 117

Qy 61 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 101

Db 118 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 158

RESULT 13

AR487264 LOCUS 470 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 286 from patent US 6706484.

ACCESSION AR487264

VERSION AR487264.1 GI:47252215

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Pluckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL Patent: US 6706484-A 286 16-MAR-2004;

FEATURES Location/Qualifiers

source 1..470

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;

Best Local Similarity 100.0%; Pred. No. 6.1e-16;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 60

Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 117

Qy 61 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 101

Db 118 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 158

RESULT 14

AR160383 LOCUS 472 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 5 from patent US 6255071.

ACCESSION AR160383

VERSION AR160383.1 GI:16224205

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 472)

AUTHORS Beach, D.H., Hannon, G.J., Conklin, D. and Sun, P.

TITLE Mammalian viral vectors and their uses

JOURNAL Patent: US 6255071-A 5 03-JUL-2001;

FEATURES Location/Qualifiers

source 1..472

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 472;

Best Local Similarity 100.0%; Pred. No. 6.1e-16;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 60

Db 63 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTTCCTTCCT 122

Qy 61 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 101

Db 123 TTCTCGCACGCTTCGCGGCTTTCCCGCTCAAGCTCTAAAT 163

RESULT 15

BD194797 LOCUS 472 bp DNA linear PAT 17-JUL-2003

DEFINITION Viral vectors and their uses.

ACCESSION BD194797

VERSION BD194797.1 GI:33004545

KEYWORDS JP 2002514054-A/6.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 472)



AUTHORS Beach,D.H., Hannon,G.J., Conklin,D.S. and Sun,P.  
TITLE Viral vectors and their uses  
JOURNAL Patent: JP 2002514054-A 6 14-MAY-2002;  
COLD SPRING HARBOR LABORATORY  
COMMENT OS Unidentified  
PN JP 2002514054-A/6  
PD 14-MAY-2002  
PF 22-SEP-1997 JP 1998515028  
PR 20-SEP-1996 US 08/716926,19-MAR-1997 US 08/820931 PI  
DAVID H BEACH,GREGORY J HANNON,DOUGLAS S CONKLIN,PEIQUING SUN PC  
C12N15/86,C12N15/10,C07K14/025  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Viral vectors and their uses  
FH Key Location/Qualifiers  
FT source 1..472  
FT /organism='Unidentified'.  
FEATURES Location/Qualifiers  
source 1..472  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 472;  
Best Local Similarity 100.0%; Pred. No. 6.1e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTTCTTCCTTCCT 60  
Db |||||||  
Qy 63 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTTCTTCCTTCCT 122  
Db |||||||  
Qy 61 TTCTGCCACGTTCCCGGCTTTCCCGCTCAAGCTCTAAAT 101  
Db |||||||  
Qy 123 TTCTGCCACGTTCCCGGCTTTCCCGCTCAAGCTCTAAAT 163  
Db |||||||

Search completed: July 14, 2005, 14:03:36  
Job time : 758.618 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_3565\_3665  
Perfect score: 101  
Sequence: 1 gcgtgaccgtacactgccc.....ttcccgctcaagctctaaat 101

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001as:.\*  
5: Geneseq2001bs:.\*  
6: Geneseq2002as:.\*  
7: Geneseq2002bs:.\*  
8: Geneseq2003as:.\*  
9: Geneseq2003bs:.\*  
10: Geneseq2003cs:.\*  
11: Geneseq2003ds:.\*  
12: Geneseq2004as:.\*  
13: Geneseq2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	288	4	AAS33391 DNA encod
2	101	100.0	288	4	AAS33392 DNA encod
3	101	100.0	288	4	AAS33392 Human imm
4	101	100.0	288	4	AAS33392 Human imm
5	101	100.0	288	4	AAS33392 Human rep
6	101	100.0	288	4	AAS33392 Human rep
7	101	100.0	288	8	ABZ74603 Secreted
8	101	100.0	288	8	ABZ74602 Secreted
9	101	100.0	288	8	ADA98980 Human sec
10	101	100.0	288	8	ADA98979 Human sec
11	101	100.0	288	8	ADA98979 Human sec
12	101	100.0	288	8	ADA98979 Human sec
13	101	100.0	288	10	ABZ68124 Human sec
14	101	100.0	288	10	ABZ68123 Human sec
15	101	100.0	288	10	ABZ68100 Human sec
16	101	100.0	288	10	ABZ68009 Human sec
17	101	100.0	299	12	ADO27428 Z. ballii
18	101	100.0	354	8	ABX38851 Bovine ES
19	101	100.0	354	8	ABX46900 Bovine ES
20	101	100.0	355	8	ABX38339 Bovine ES

21	101	100.0	359	8	ABX38602 Bovine ES
22	101	100.0	359	8	ABX41939 Bovine ES
23	101	100.0	359	8	ABX42410 Bovine ES
24	101	100.0	361	8	ABX46893 Bovine ES
25	101	100.0	361	8	ABX48545 Bovine ES
26	101	100.0	362	8	ABX46167 Bovine ES
27	101	100.0	362	8	ABX38860 Bovine ES
28	101	100.0	364	8	ABX46417 Bovine ES
29	101	100.0	365	8	ABX39110 Bovine ES
30	101	100.0	365	8	ABX41262 Bovine ES
31	101	100.0	365	8	ABX46903 Bovine ES
32	101	100.0	365	8	ABX38856 Bovine ES
33	101	100.0	366	8	ABX39785 Bovine ES
34	101	100.0	366	8	ABX43577 Bovine ES
35	101	100.0	366	8	ABX43580 Bovine ES
36	101	100.0	366	8	ABX45460 Bovine ES
37	101	100.0	366	8	ABX44973 Bovine ES
38	101	100.0	366	8	ABX48549 Bovine ES
39	101	100.0	366	8	ABX35989 Bovine ES
40	101	100.0	367	8	ABX37877 Bovine ES
41	101	100.0	367	8	ABX39577 Bovine ES
42	101	100.0	368	8	ABX42411 Bovine ES
43	101	100.0	368	8	ABX35306 Bovine ES
44	101	100.0	369	8	ABX36699 Bovine ES
45	101	100.0	369	8	ABX40761 Bovine ES

ALIGNMENTS

RESULT 1  
AAS33391  
ID AAS33391 standard; DNA; 288 BP.

AC AAS33391;

XX 04-DEC-2001 (first entry)

DE DNA encoding human secreted protein, Seq ID No 674.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
XX rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;  
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
XX cystostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
XX multiple sclerosis; cancer; hyperproliferative disorder; infection;  
XX Gaucher's disease; neurological disease; cerebrovascular disorder;  
XX thrombosis; wound healing; ds.

OS Homo sapiens.

PN WO200155326-A2.

XX 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216800P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.



CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and  
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular  
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists  
CC and antibodies can also be used to promote wound healing, maintain organs  
CC before transplantation, and support cell culture of primary tissues. PCR  
CC AAS33043-AAS33486 represent human secreted protein coding sequences. PCR

Query Match 100.0%; Score 101; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCCTTCCTTCCTTCCT 60  
Db 33 CGGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCCTTCCTTCCTTCCT 92  
Qy 61 TTCTCGCCACGCTTCGCGCGCTTCCCGTCAAGCTCTAAAT 101  
Db 93 TTCTCGCCACGCTTCGCGCGCTTCCCGTCAAGCTCTAAAT 133

## RESULT 2

AAS333392 standard; DNA; 288 BP.

AC AAS333392;

DT 04-DEC-2001 (first entry)

DE DNA encoding human secreted protein, Seq ID No 675.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cystostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ds.

XX Homo sapiens.

XX WO200155326-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225577P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-451931/48.  
 XX  
 XX New nucleic acids and polypeptides, useful for diagnosing, preventing or  
 PT treating medical conditions.  
 XX  
 XX Disclosure; SEQ ID NO 675; 753pp; English.  
 XX  
 XX The invention relates to novel isolated nucleic acid molecules (I)  
 CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
 CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
 CC the prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate expression of secreted proteins. (I) and complementary  
 CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
 CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
 CC of similar nucleic acid sequences in samples, and so which patients may  
 CC be in need of restorative therapy. (II) may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of the secreted  
 CC proteins. The anti-(II) antibodies and antagonists may also be used to  
 CC down regulate expression and activity of (II). The anti-(II) antibodies  
 CC may also be used as diagnostic agents for detecting the presence of (II)  
 CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The  
 CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis

CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
 CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and  
 CC thrombosis), infections caused by bacteria, viruses and fungi and ocular  
 CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists  
 CC and antibodies can also be used to promote wound healing, maintain organs  
 CC and before transplantation, and support cell culture of primary tissues.  
 CC AAS33043-AAS33486 represent human secreted protein coding sequences, PCR

Query Match 100.0%; Score 101; DB 4; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-21; Mismatches 0; Gaps 0;  
 Matches 101; Conservative 0; Indels 0; Gaps 0;

QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGTTTCTTCCCTCCT 60  
 DB 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGTTTCTTCCCTCCT 92  
 QY 61 TTCTCGCCAGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101  
 DB 93 TTCTCGCCAGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133

RESULT 3  
 AAK87482  
 ID AAK87482 standard; DNA; 288 BP.  
 XX  
 AC AAK87482;  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42294.  
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis; ds.  
 XX Homo sapiens.  
 XX  
 PN WO200157182-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 XX 17-JAN-2001; 2001WO-US001354.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246174P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rojen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.

XX PS Disclosure; SEQ ID NO 42294; 3071pp + Sequence Listing; English.

XX CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention

XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 288;

Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTGACCGGTACACTTGGCCAGCGCCCTAGCGCGCTTCCTTCCTTCCTTCCT 60  
|||||

Db 33 GCGTCACCGCTACACTTCCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCCTTCCT 92  
Qy 61 TTCTCGCACGTTTCGCGGCTTTCGCGTCAAGCTCTAAAT 101  
Db 93 TTCTCGCACGTTTCGCGGCTTTCGCGTCAAGCTCTAAAT 133

RESULT 4  
AAK87481  
ID AAK87481 standard; DNA; 288 BP.  
XX AC AAK87481;  
XX AC  
XX AC  
DT 07-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42293.  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX WO200157182-A2.  
XX  
XX  
XX 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-024617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246533P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.



PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Disclosure; SEQ ID NO 42293; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 101; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTCTTCCTTCCT 60  
Db 33 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTCTTCCTTCCT 92  
QY 61 TTCTCGCACGTTCCCGGCTTCCCGCTCAAGCTCTAAAT 101  
Db 93 TTCTCGCACGTTCCCGGCTTCCCGCTCAAGCTCTAAAT 133  
RESULT 5  
ID AAL07028 standard; DNA; 288 BP.  
XX  
XX AAL07028;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 9716.  
DE  
XX

KW Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ds.  
XX Homo sapiens.  
OS  
XX WO200155320-A2.  
FN  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US001339.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226273P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251898P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX Disclosure; SEQ ID NO 9716; 1297pp + Sequence Listing; English.
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCGCTAGCGCCGCTTCCCGCTCAAGCTCTAAAT 101
Db 33 GCGTGACCGCTACACTTGCAGCGCGCTAGCGCCGCTTCCCGCTTCTTCCCTTCCCT 60
Qy 61 TTCTCGCCAGTTTCGGCGGCTTTCCCGCTCAAGCTCTAAAT 101
Db 93 TTCTCGCCAGTTTCGGCGGCTTTCCCGCTCAAGCTCTAAAT 133

RESULT 6
AAL07029
ID AAL07029 standard; DNA; 288 BP.
XX
AC AAL07029;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 9717.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
```



QY 61 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101  
 |||||  
 Db 93 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133  
 |||||

RESULT 7  
 ABZ74603  
 ID ABZ74603 standard; DNA; 288 BP.  
 XX  
 AC ABZ74603;  
 XX  
 DT 12-MAY-2003 (first entry)  
 XX  
 DE Secreted protein gene 366 genomic fragment HUSGU40, SEQ ID NO:1750.  
 XX  
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;  
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
 KW drug screening; chromosome identification; chromosome mapping;  
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;  
 KW antianaemic; vulnery; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200277013-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 XX 26-MAR-2002; 2002WO-US009370.  
 XX  
 XX 27-MAR-2001; 2001US-0278650P.  
 XX  
 PR 12-SEP-2001; 2001US-00950082.  
 PR  
 PR 12-SEP-2001; 2001US-00950083.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 XX  
 XX WPI; 2003-040578/03.  
 DR  
 XX New human secreted proteins and nucleic acids, useful for detecting or  
 PT treating cancer or other hyperproliferative disorders, autoimmune  
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.  
 XX  
 PS Disclosure; Page 2326; 2474pp; English.

CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
 CC protein genes, and ABP0947-ABP01363 represent the proteins they encode.  
 CC CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins are thought to be involved in biological activities  
 CC associated with cellular signalling, cellular differentiation, cell  
 CC migration, prohormone activation and neurotransmitter activity. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing or treating cancers or other  
 CC hyperproliferative disorders. Additionally, the secreted proteins and  
 CC their nucleic acids may also be used in the treatment of autoimmune  
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS  
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote  
 CC wound healing. Nucleic acids of the invention may be used for chromosome  
 CC identification, chromosome mapping, in gene therapy, for identifying  
 CC individuals from minute biological samples, as hybridisation probes, and  
 CC as molecular weight markers. The present sequence represents a human  
 CC secreted protein genomic fragment referred to in the disclosure of the  
 CC invention

Seq Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 8; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGACCGCTACACTTGCAGCGCCTAGCGCCGCTCCTTCGCTTTCTTCCTTCCT 60  
 |||||  
 Db 33 GCGTGACCGCTACACTTGCAGCGCCTAGCGCCGCTCCTTCGCTTTCTTCCTTCCT 92  
 |||||

QY 61 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101  
 |||||  
 Db 93 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133  
 |||||

RESULT 8  
 ABZ74602  
 ID ABZ74602 standard; DNA; 288 BP.  
 XX  
 AC ABZ74602;  
 XX  
 DT 12-MAY-2003 (first entry)  
 XX  
 DE Secreted protein gene 366 genomic fragment HUSGU40, SEQ ID NO:1749.  
 XX  
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;  
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
 KW drug screening; chromosome identification; chromosome mapping;  
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;  
 KW antianaemic; vulnery; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200277013-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 XX 26-MAR-2002; 2002WO-US009370.  
 XX  
 XX 27-MAR-2001; 2001US-0278650P.  
 XX  
 PR 12-SEP-2001; 2001US-00950082.  
 PR  
 PR 12-SEP-2001; 2001US-00950083.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 XX  
 XX WPI; 2003-040578/03.  
 DR  
 XX New human secreted proteins and nucleic acids, useful for detecting or  
 PT treating cancer or other hyperproliferative disorders, autoimmune  
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.  
 XX  
 PS Disclosure; Page 2326; 2474pp; English.

CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
 CC protein genes, and ABP0947-ABP01363 represent the proteins they encode.  
 CC CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins are thought to be involved in biological activities  
 CC associated with cellular signalling, cellular differentiation, cell  
 CC migration, prohormone activation and neurotransmitter activity. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing or treating cancers or other  
 CC hyperproliferative disorders. Additionally, the secreted proteins and  
 CC their nucleic acids may also be used in the treatment of autoimmune  
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS  
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote  
 CC wound healing. Nucleic acids of the invention may be used for chromosome  
 CC identification, chromosome mapping, in gene therapy, for identifying  
 CC individuals from minute biological samples, as hybridisation probes, and  
 CC as molecular weight markers. The present sequence represents a human  
 CC secreted protein genomic fragment referred to in the disclosure of the  
 CC invention

```
CC invention
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match      100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCAAGCTCTAAAT 101
Db 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCTTCCCTTCCCT 60
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133

RESULT 9
ADA98980
ID ADA98980 standard; DNA; 288 BP.
AC ADA98980;
XX
XX
DT 20-NOV-2003 (first entry)
XX
XX Human secreted protein-related DNA sequence #573.
XX human; secreted protein; cardiovascular disorder; arrhythmia;
XX atherosclerosis; stroke; endocarditis; congestive heart failure;
XX rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
XX migraine; thrombosis; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
XX Homo sapiens.
XX OS
XX PN WO2003004623-A2.
XX FN
XX PD 16-JAN-2003.
XX PF
XX PP 26-MAR-2002; 2002WO-US009922.
XX PR 27-MAR-2001; 2001US-0278650P.
XX PR 12-SEP-2001; 2001US-00950082.
XX PR 12-SEP-2001; 2001US-00950083.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2003-247946/24.
XX
XX New human secreted polypeptide and nucleic acid molecules, useful for
XX diagnosing, preventing, prognosticating or treating cardiovascular
XX disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
XX thrombosis).
XX
XX Disclosure; SEQ ID NO 1089; 1572pp; English.
XX
XX The invention comprises the amino acid and coding sequence of human
XX secreted proteins. The DNA and protein sequences of the invention are
XX useful in the treatment of cardiovascular disorders, such as: arrhythmia,
XX atherosclerosis, stroke, endocarditis, congestive heart failure,
XX rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
XX migraine, or thrombosis. The DNA and protein sequences may also be used
XX for treating or preventing: neural disorders, immune system disorders,
XX muscular disorders, reproductive disorders, gastrointestinal disorders,
XX pulmonary disorders, renal disorders, proliferative disorders and/or
XX cancerous diseases. The present DNA sequence is used in the
XX exemplification of the invention. NOTE: The present sequence is shown on
XX the WIPO website.
XX
XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match      100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCAAGCTCTAAAT 101
Db 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCTTCCCTTCCCT 60
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133

RESULT 10
ADA98979
ID ADA98979 standard; DNA; 288 BP.
XX
XX AC ADA98979;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Human secreted protein-related DNA sequence #572.
XX
XX KW human; secreted protein; cardiovascular disorder; arrhythmia;
XX atherosclerosis; stroke; endocarditis; congestive heart failure;
XX KW rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
XX KW migraine; thrombosis; neural disorder; immune system disorder;
XX KW muscular disorder; reproductive disorder; gastrointestinal disorder;
XX KW pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
XX OS Homo sapiens.
XX
XX PN WO2003004623-A2.
XX
XX PD 16-JAN-2003.
XX
XX PF 26-MAR-2002; 2002WO-US009922.
XX
XX PR 27-MAR-2001; 2001US-0278650P.
XX
XX PR 12-SEP-2001; 2001US-00950082.
XX
XX PR 12-SEP-2001; 2001US-00950083.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Ruben SM;
XX
XX DR WPI; 2003-247946/24.
XX
XX New human secreted polypeptide and nucleic acid molecules, useful for
XX diagnosing, preventing, prognosticating or treating cardiovascular
XX disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
XX thrombosis).
XX
XX Disclosure; SEQ ID NO 1088; 1572pp; English.
XX
XX The invention comprises the amino acid and coding sequence of human
XX secreted proteins. The DNA and protein sequences of the invention are
XX useful in the treatment of cardiovascular disorders, such as: arrhythmia,
XX atherosclerosis, stroke, endocarditis, congestive heart failure,
XX KW rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
XX KW migraine, or thrombosis. The DNA and protein sequences may also be used
XX for treating or preventing: neural disorders, immune system disorders,
XX muscular disorders, reproductive disorders, gastrointestinal disorders,
XX pulmonary disorders, renal disorders, proliferative disorders and/or
XX cancerous diseases. The present DNA sequence is used in the
XX exemplification of the invention. NOTE: The present sequence is shown on
XX the WIPO website.
XX
XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match      100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCAAGCTCTAAAT 101
Db 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGCCCTTCCCGTCTTCCCTTCCCT 60
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133
```

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 60  
Db |||||  
33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 92  
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 101  
Db |||||  
93 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 133

## RESULT 11

ADA44489  
ID ADA44489 standard; DNA; 288 BP.

XX AC ADA44489;

XX DT 20-NOV-2003 (first entry)

XX DE Human secreted protein DNA SEQ ID 682.

XX KW Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;  
XX KW Neuroprotective; Cerebroprotective; Antianemic; ds.  
XX OS Homo sapiens.  
XX PN WO2003009865-A2.  
XX PD 03-JAN-2003.  
XX PF 26-MAR-2002; 2002WO-US009105.  
XX PR 27-MAR-2001; 2001US-0278650P.  
XX PR 12-SEP-2001; 2001US-00950082.  
XX PR 12-SEP-2001; 2001US-00950083.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM;  
XX WPI; 2003-184045/18.

XX PT A human secreted protein and nucleic acids useful for preparing a  
diagnostic or pharmaceutical composition for diagnosing or treating  
diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,  
retinopathy, neuropathy.  
XX PS Disclosure; SEQ ID NO 682; 701pp; English.  
XX CC The invention relates to novel genes and their fragments which are useful  
for preventing, treating or ameliorating medical conditions e.g. by  
protein or gene therapy. The genes are isolated from a range of human  
tissues disclosed in the specification. The nucleic acids and proteins  
are useful in the diagnosis, treatment and prevention of conditions  
related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,  
polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,  
infection, cataract, renal disorders, or endocrine disorders. The present  
sequence was used to illustrate the invention.

XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

XX CC Query Match 100.0%; Score 101; DB 8; Length 288;  
XX CC Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
XX CC Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX CC 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 60  
XX CC |||||  
XX CC 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 92  
XX CC 61 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 101  
XX CC |||||  
XX CC 93 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 133

XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

XX CC Query Match 100.0%; Score 101; DB 8; Length 288;  
XX CC Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
XX CC Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 60  
Db |||||  
33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 92

Qy 61 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 101

Db |||||

93 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 133

## RESULT 12

ADA44488  
ID ADA44488 standard; DNA; 288 BP.

XX AC ADA44488;

XX DT 20-NOV-2003 (first entry)

XX DE Human secreted protein DNA SEQ ID 681.

XX KW Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;  
XX KW Neuroprotective; Cerebroprotective; Antianemic; ds.  
XX OS Homo sapiens.  
XX PN WO2003009865-A2.  
XX PD 03-JAN-2003.  
XX PF 26-MAR-2002; 2002WO-US009105.  
XX PR 27-MAR-2001; 2001US-0278650P.  
XX PR 12-SEP-2001; 2001US-00950082.  
XX PR 12-SEP-2001; 2001US-00950083.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM;  
XX WPI; 2003-184045/18.

XX PT A human secreted protein and nucleic acids useful for preparing a  
diagnostic or pharmaceutical composition for diagnosing or treating  
diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,  
retinopathy, neuropathy.  
XX PS Disclosure; SEQ ID NO 681; 701pp; English.  
XX CC The invention relates to novel genes and their fragments which are useful  
for preventing, treating or ameliorating medical conditions e.g. by  
protein or gene therapy. The genes are isolated from a range of human  
tissues disclosed in the specification. The nucleic acids and proteins  
are useful in the diagnosis, treatment and prevention of conditions  
related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,  
polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,  
infection, cataract, renal disorders, or endocrine disorders. The present  
sequence was used to illustrate the invention.

XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

XX CC Query Match 100.0%; Score 101; DB 8; Length 288;  
XX CC Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
XX CC Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX CC 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 60  
XX CC |||||  
XX CC 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 92  
XX CC 61 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 101  
XX CC |||||  
XX CC 93 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 133

XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

XX CC Query Match 100.0%; Score 101; DB 8; Length 288;  
XX CC Best Local Similarity 100.0%; Pred. No. 4.4e-21;  
XX CC Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 60  
Db |||||  
33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCTTCCCT 92

Qy 61 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 101

Db |||||

93 TTCTCGCACGTTCCGCGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 133

## RESULT 13

ABZ68124  
ID ABZ68124 standard; DNA; 288 BP.

XX AC ABZ68124;

XX DT 26-MAR-2003 (first entry)

XX DE Human secreted protein encoding genomic DNA SEQ ID NO 1647.

```
XX Human; secreted protein; neutrotropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antiskickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX Homo sapiens.
XX WO200277186-A2.
XX 03-OCT-2002.
XX 26-MAR-2002; 2002WO-US009188.
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX Disclosure; Page 2268; 2423pp; English.
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTCCTTCCTTCCT 60
DB 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTCCTTCCTTCCT 92
QY 61 TTCTCGCCACGTTGCGCGGCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 101
DB 93 TTCTCGCCACGTTGCGCGGCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 133
RESULT 14
ID ABZ68123
XX AC ABZ68123;
XX 26-MAR-2003 (first entry)
DT
```

```
XX Human secreted protein encoding genomic DNA SEQ ID NO 1646.
XX Human; secreted protein; neutrotropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antiskickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX Homo sapiens.
XX WO200277186-A2.
XX 03-OCT-2002.
XX 26-MAR-2002; 2002WO-US009188.
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX Disclosure; Page 2268; 2423pp; English.
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTCCTTCCTTCCT 60
DB 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTCCTTCCTTCCT 92
QY 61 TTCTCGCCACGTTGCGCGGCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 101
DB 93 TTCTCGCCACGTTGCGCGGCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 133
RESULT 15
ID ABZ68010
XX AC ABZ68010;
XX 26-MAR-2003 (first entry)
DT
```

XX	26-MAR-2003	(first entry)
DT	Human secreted protein encoding genomic DNA SEQ ID NO 1533.	
XX		
DE		
XX		
XX	Human; secreted protein; neutropic; neuroprotective; cytostatic; virucide; dermatological; immunosuppressive; antinflammatory; anti-HIV; KW	
KW	vulnerable; antibacterial; antiparkinsonian; antickling; antianaemic; KW	
KW	antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; KW	
KW	antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant; KW	
KW	antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine; KW	
KW	cardiovascular disorder; neurological disease; nephrotropic; KW	
KW	gene therapy; gene; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	W0200277186-A2.	
XX	03-OCT-2002.	
PD		
XX		
PF	26-MAR-2002; 2002WO-US009188.	
XX		
XX	27-MAR-2001; 2001US-0278650P.	
PR	12-SEP-2001; 2001US-00950082.	
PR	12-SEP-2001; 2001US-00950083.	
XX		
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA		
XX		
PI	Rosen CA, Ruben SM;	
XX		
XX	WPI; 2003-040583/03.	
DR		
XX		
PT	New human secreted proteins encoded by genes contained in cDNA clones (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS, PT	
PT	multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or PT	
PT	West Nile fever.	
XX		
PS	Disclosure; Page 2154; 2423pp; English.	
XX		
CC	The invention relates to novel human genes (ABZ66891-ABZ68209) and the CC	
CC	encoded secreted proteins (ABP99470-ABP99872) useful for preventing, CC	
CC	treating or ameliorating medical conditions e.g. by protein or gene CC	
CC	therapy. The genes are isolated from a range of human tissues disclosed CC	
CC	in the specification. The nucleic acids, proteins, antibodies and CC	
CC	(antagonists are useful in the diagnosis, treatment and prevention of: CC	
CC	(a) cancer, e.g. breast and ovarian cancer and other cancers of the CC	
CC	adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, CC	
CC	lung or urogenital; (b) immune disorders e.g. Addison's disease, CC	
CC	allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, CC	
CC	diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid CC	
CC	arthritis and ulcerative colitis; (c) cardiovascular disorders such as CC	
CC	myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. CC	
CC	cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, CC	
CC	bacterial, fungal and parasitic infections	
XX		
SQ	Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;	
Query Match 100.0%; Score 101; DB 10; Length 288;		
Best Local Similarity 100.0%; Pred. No. 4.4e-21;		
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 GCGTGACCGCTACACTTGGCAGCGCCCTAGCGCCGCTCTCTTCGCTTTCTTCCTTCCT 60	
Db	33 GCGTGACCGCTACACTTGGCAGCGCCCTAGCGCCGCTCTCTTCGCTTTCTTCCTTCCT 92	
Qy	61 TTCTCGGCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101	
Db	93 TTCTCGGCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_3565\_3665

Perfect score: 101

Sequence: 1 gcgtgaccgtacacttgcc.....ttccccgcgaagtctctaaat 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	113	8	AQ265696
2	101	100.0	116	8	AQ263357
3	101	100.0	142	8	AQ265335
4	101	100.0	150	8	AQ264698
5	101	100.0	158	8	AQ026683
6	101	100.0	162	8	AQ061259
7	101	100.0	163	8	AQ196514
8	101	100.0	164	4	BW812626
9	101	100.0	166	8	AQ112846
10	101	100.0	168	4	BW812644
11	101	100.0	173	8	AQ310689
12	101	100.0	192	8	AQ280364
13	101	100.0	200	8	AQ278442
14	101	100.0	211	8	AQ280347
15	101	100.0	217	4	B1677411
16	101	100.0	218	8	AQ280220
17	101	100.0	224	4	BW517095
18	101	100.0	230	8	AQ279924
19	101	100.0	236	6	CD282304
20	101	100.0	249	4	B1677389
21	101	100.0	255	4	B1940501
22	101	100.0	261	8	AQ310638
23	101	100.0	264	6	CD280795
24	101	100.0	271	8	B83638

25	101	100.0	271	8	AQ427025
26	101	100.0	275	8	AZ212954
27	101	100.0	276	6	CB867747
28	101	100.0	277	8	AQ263549
29	101	100.0	281	2	BF542624
30	101	100.0	286	7	CK884250
31	101	100.0	286	8	AQ074694
32	101	100.0	288	5	BW573920
33	101	100.0	292	8	B78686
34	101	100.0	307	8	B78705
35	101	100.0	309	7	CK886580
36	101	100.0	309	7	CV162119
37	101	100.0	310	8	AQ009663
38	101	100.0	312	2	BF524055
39	101	100.0	324	8	AQ008830
40	101	100.0	324	8	AQ017932
41	101	100.0	325	6	CD279858
42	101	100.0	325	6	CD281090
43	101	100.0	327	6	CD282801
44	101	100.0	327	8	AQ026530
45	101	100.0	332	8	AQ009332

#### ALIGNMENTS

RESULT 1  
LOCUS AQ265696 113 bp DNA linear GSS 27-OCT-1998  
DEFINITION CITBI-El-2503El.TR CITBI-El Homo sapiens genomic clone 2503El,  
genomic survey sequence.

ACCESSION AQ265696

VERSION AQ265696.1 GI:3791450

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 113)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Adams M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,  
Barry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and  
Venter, J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready

JOURNAL Map Building

COMMENT Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.

Seg primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1. .113

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="2503El"

/sex="male"

/cell\_type="sperm"

/clone\_lib="CITBI-El"

/note="Vector: pBelOBAcl1; Site\_1: EcoRI; Site\_2: EcoRI;

CalTech Human BAC Library D"

ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 113;

Best Local Similarity 100.0%; Pred. No. 9.1e-18;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 2
AQ263357 116 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2503C21.TR CITBI-EI Homo sapiens genomic clone 2503C21,
DEFINITION genomic survey sequence.
ACCESSION AQ263357
VERSION AQ263357.1 GI:3790953
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 116)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source
     1..116
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /clone="2503C21"
     /sex="male"
     /cell_type="sperm"
     /clone_lib="CITBI-EI"
     /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 3
AQ265335 142 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2508M1.TR CITBI-EI Homo sapiens genomic clone 2508M1,
DEFINITION genomic survey sequence.

```

```

ACCESSION AQ265335
VERSION AQ265335.1 GI:3793535
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source
     1..142
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /clone="2508M1"
     /sex="male"
     /cell_type="sperm"
     /clone_lib="CITBI-EI"
     /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 142;
Best Local Similarity 100.0%; Pred. No. 9.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 4
AQ264698 150 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2508E1.TR CITBI-EI Homo sapiens genomic clone 2508E1,
DEFINITION genomic survey sequence.
ACCESSION AQ264698
VERSION AQ264698.1 GI:3792898
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 150)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics

```

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
Seq primer: M13 Reverse  
Class: BAC ends.

#### FEATURES

source

Location/Qualifiers  
1. .150  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2508E1"  
/sex="male"  
/cell\_type="sperm"  
/clone\_lib="CITBi-E1"  
/note="Vector: pBelobAC11; Site 1: EcoRI; Site 2: EcoRI;  
CalTech Human BAC Library D"

#### ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 150;  
Best Local Similarity 100.0%; Pred. No. 9.3e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60  
Db 8 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 67  
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 68 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 108

#### RESULT 5

AQ026683  
LOCUS AQ026683 158 bp DNA linear GSS 30-JUN-1998  
DEFINITION CIT-HSP-2314E2.TR CIT-HSP Homo sapiens genomic clone 2314E2,  
genomic survey sequence.

ACCESSION AQ026683

VERSION AQ026683.1 GI:3266905

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 158)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).

Seq primer: M13 Reverse

Class: BAC ends.

#### FEATURES

source

Location/Qualifiers  
1. .158  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2314E2"

/sex="Male"

/cell\_type="Sperm"

/clone\_lib="CIT-HSP"

/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

#### ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 158;  
Best Local Similarity 100.0%; Pred. No. 9.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60  
Db 52 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 111  
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101  
Db 112 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 152

#### RESULT 6

AQ061259 162 bp DNA linear GSS 31-JUL-1998  
LOCUS AQ061259  
DEFINITION CIT-HSP-2352N1.TR CIT-HSP Homo sapiens genomic clone 2352N1,  
genomic survey sequence.

ACCESSION AQ061259

VERSION AQ061259.1 GI:3363171

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 162)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).

Seq primer: M13 Reverse

Class: BAC ends.

#### FEATURES

source

Location/Qualifiers  
1. .162  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2352N1"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

#### ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 162;  
Best Local Similarity 100.0%; Pred. No. 9.4e-18;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60  
Db 52 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 111  
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101

```

Db      112 TTCTCGCACGTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 152
|||||
RESULT 7
AQ196514
LOCUS
DEFINITION
CIT-HSP-2385C23.TR CIT-HSP Homo sapiens genomic clone 2385C23,
163 bp DNA linear GSS 16-SEP-1998
genomic survey sequence.
ACCESSION
AQ196514
VERSION
AQ196514.1 GI:3603876
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 163)
AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL
Unpublished (1998)
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1..163
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2385C23"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 163;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCCCTTCCT 60
|||||
Db 52 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCCCTTCCT 111
|||||
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGCTCAAGCTCTAAAT 101
|||||
Db 112 TTCTCGCACGTTCCGCGGCTTTCCCGCTCAAGCTCTAAAT 152
|||||
RESULT 8
BM812626
LOCUS
DEFINITION
rt03b06.y2 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy
Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION
BM812626
VERSION
BM812626.1 GI:19148640
KEYWORDS
EST.
SOURCE
Pristionchus pacificus
ORGANISM
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
1 (bases 1 to 164)
AUTHORS
McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Bennett,J., Franklin,B.,
Bowers,Y., Gibbons,M., Ritter,E., Kennet,J., Maguire,L., Beck,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE
The Washington Univ. Nematode EST Project, 1999
JOURNAL
Unpublished (1999)
COMMENT
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli, and
Dr. James McCarter at Washington University, St. Louis. DNA
Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40RP from Gibco
High quality sequence stop: 161.
FEATURES
source
1..164
/organism="Pristionchus pacificus"
/mol_type="mRNA"
/db_xref="taxon:54126"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
Murphy Chiapelli McCarter"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
Site_2: EcoRI; The library was constructed by Claire
Murphy, Brandi Chiapelli, and Dr. James McCarter at
Washington University, St. Louis. Oligo(dT)-SL1 PCR based
library. Pristionchus pacificus mixed stage cDNA PCR
products of size 3400 nucleotides containing SL1 on the
5' end and oligo(dT) on the 3' end were non-directionally
cloned into pCRII-TOPO(Invitrogen) following the TOPO TA
cloning protocol."
ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCCCTTCCT 60
|||||
Db 51 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTTTCGCTTCTTCCCTTCCT 110
|||||
Qy 61 TTCTCGCACGTTCCGCGGCTTTCCCGCTCAAGCTCTAAAT 101
|||||
Db 111 TTCTCGCACGTTCCGCGGCTTTCCCGCTCAAGCTCTAAAT 151
|||||
RESULT 9
AQ112846
LOCUS
DEFINITION
CIT-HSP-2376E2.TR CIT-HSP Homo sapiens genomic clone 2376E2,
genomic survey sequence.
ACCESSION
AQ112846
VERSION
AQ112846.1 GI:3488967
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 166)
AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building

```



```

/clone_lib="CITBI-E1"
/notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 173;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 12
LOCUS      AQ280364
DEFINITION CITBI-E1-2518K13.TR CITBI-E1 Homo sapiens genomic clone 2518K13,
            genomic survey sequence.
ACCESSION  AQ280364
VERSION    AQ280364.1 GI:3906183
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 192)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,W. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..192
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2518K13"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

FEATURES             Location/Qualifiers
     source           1..192
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2518K13"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 192;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 13
LOCUS      AQ278442
DEFINITION CITBI-E1-2519N1.TR CITBI-E1 Homo sapiens genomic clone 2519N1,
            genomic survey sequence.
ACCESSION  AQ278442
VERSION    AQ278442.1 GI:3904410
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 200)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,W. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..200
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2519N1"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 200;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 12 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 71

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 72 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 112

RESULT 14
LOCUS      AQ280347
DEFINITION CITBI-E1-2518K3.TR CITBI-E1 Homo sapiens genomic clone 2518K3,
            genomic survey sequence.
ACCESSION  AQ280347
VERSION    AQ280347.1 GI:3906166
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 211)
```

```

68 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 13
LOCUS      AQ278442
DEFINITION CITBI-E1-2519N1.TR CITBI-E1 Homo sapiens genomic clone 2519N1,
            genomic survey sequence.
ACCESSION  AQ278442
VERSION    AQ278442.1 GI:3904410
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 200)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,W. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..200
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2519N1"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 200;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 60
Db 12 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTTCGCTTCTTCCCTTCCT 71

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 72 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 112

RESULT 14
LOCUS      AQ280347
DEFINITION CITBI-E1-2518K3.TR CITBI-E1 Homo sapiens genomic clone 2518K3,
            genomic survey sequence.
ACCESSION  AQ280347
VERSION    AQ280347.1 GI:3906166
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 211)
```

**AUTHORS** Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.  
**TITLE** Use of a random human BAC End Sequence Database for Sequence-Ready  
**JOURNAL** Map Building  
**COMMENT** Unpublished (1998)  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/tldb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: M13 Reverse  
 Class: BAC ends.

#### FEATURES

source  
 1. .211  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2518K3"  
 /sex="male"  
 /cell\_type="sperm"  
 /clone\_lib="CITBI-E1"  
 /note="Vector: pBelobAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
 Caltech Human BAC Library D"

#### ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 60  
 Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 67  
 Qy 61 TTCTCGCCACGTTCCCGCGCTTTCCTCGCTCAAGCTCTAAAT 101  
 Db 68 TTCTCGCCACGTTCCCGCGCTTTCCTCGCTCAAGCTCTAAAT 108

RESULT 15  
 BI677411  
 LOCUS BI677411 217 bp mRNA linear EST 17-SEP-2001  
 DEFINITION ie30903.y1 Kaestner ngn3 wt Mus musculus cDNA 5', mRNA sequence.  
 ACCESSION BI677411  
 VERSION BI677411.1 GI:15630318  
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 217)  
 Melton,D., Brown,J., Kenty,G., Brestelli,J., Gradwohl,G., Clifton,S.,  
 Lenisha,I., Searce,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,  
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,  
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,  
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,  
 Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)

**TITLE** Endocrine Pancreas Consortium  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohph.harvard.edu  
 Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,

2000) Library was constructed by Catherine Lee DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Marie Searce  
 (mscearc@mail.med.upenn.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 207.  
 Location/Qualifiers  
 1. .217  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129/Sv x CD1"  
 /db\_xref="taxon:10090"  
 /dev\_stage="p.c. 14.5"  
 /lab\_host="E. coli-DH12S (GIBCO)"  
 /clone\_lib="Kaestner ngn3 wt"  
 /note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site\_1:  
 Not I; Site\_2: Sal I; The library was prepared by  
 Catherine S. Lee and has not been published. The pancreas  
 was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,  
 2000). The cDNA's were prepared with an oligo containing a  
 NotI site, and SalI linkers were added to the ends. The  
 inserts were cut with NotI before being cloned into the  
 NotI-SalI sites in the vectors. This is one of two  
 libraries, ngn3 wt and ngn3 -/- . The wt library is in  
 pSPORT1, T7 promoter is 5'."

#### ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 60  
 Db 28 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 87  
 Qy 61 TTCTCGCCACGTTCCCGCGCTTTCCTCGCTCAAGCTCTAAAT 101  
 Db 88 TTCTCGCCACGTTCCCGCGCTTTCCTCGCTCAAGCTCTAAAT 128

Search completed: July 14, 2005, 23:23:19  
 Job time : 966.667 secs

**THIS PAGE BLANK (USPTO)**



Result No.	Query %		DB	ID	Description
	Match	Score			
C 1	101	100.0	6	AR356490	AR356490 Sequence
C 2	101	100.0	142	AR538046	AR538046 Sequence
C 3	101	100.0	228	E00019	E00019 DNA coding
C 4	101	100.0	240	PMOENDO	M10199 Plasmid pMM
C 5	101	100.0	251	E00018	E00018 DNA coding
C 6	101	100.0	251	E01644	E00018 DNA coding
C 7	101	100.0	344	HUMUT5345	I01644 Sequence 1
C 8	101	100.0	400	6	Li18624 Human chr10
C 9	101	100.0	456	BD195256	BD195256 Nucleotid
C 10	101	100.0	456	E00892	E00892 Synthetic D
C 11	101	100.0	456	E01156	E01156 DNA fragmen
C 12	101	100.0	456	E01274	E01274 DNA encodin
C 13	101	100.0	456	E01302	E01302 DNA encodin
C 14	101	100.0	466	AX260098	AX260098 Sequence
C 15	101	100.0	573	AX260150	AX260150 Sequence
C 16	101	100.0	693	A43586	A43586 Sequence 11
C 17	101	100.0	998	AR116755	AR116755 Sequence
C 18	101	100.0	1011	AY559171	AY559171 Pseudomon
C 19	101	100.0	1012	SMTEMAGE	X97254 S.marcescen
C 20	101	100.0	1012	CEC11F10	Z92776 Caenorhabdi

RESULT 2					
AR538046/c	AR538046	Sequence	2608 from patent US 6737248.	142 bp	DNA linear PAT 08-OCT-2004
LOCUS	AR538046	DEFINITION			
ACCESSION	AR538046	KEYWORDS			
VERSION	AR538046.1	SOURCE	GI:53929263		
KEYWORDS	.				
SOURCE	Unknown.				

```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
source 1..142
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 60
Db 107 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 48

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 7

RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 228)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
TITLE Patent: JP 1981154999-A 2 30-NOV-1981;
JOURNAL UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00.C07H21/00.C12N1/00.C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
source 1..228
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 60
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 116

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
```

```
Db 115 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 75

RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences: plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES Location/Qualifiers
source 1..240
/organism="Plasmid pMM110"
/mol_type="genomic DNA"
/db_xref="taxon:2599"
/plasmid="Plasmid pMM110"

ORIGIN Unreported.

Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 60
Db 151 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAAACAATAG 92

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 51

RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018.1 GI:2168326
VERSION JP 1981154999-A/1.
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
1 (bases 1 to 251)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
TITLE Patent: JP 1981154999-A 1 30-NOV-1981;
JOURNAL UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00.C07H21/00.C12N1/00.C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;
```

```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
Location/Qualifiers
FEATURES
source 1..251
/organism="Escherichia coli"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
Db |||||||
175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 116
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||||
115 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 75
RESULT 6
101644/c
LOCUS 101644 251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION 101644
VERSION 101644.1 GI:267685
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 251)
AUTHORS Gilbert,W. and Talmadge,K.
TITLE Mature protein synthesis
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
President and Fellows of Harvard College; Cambridge, MA
FEATURES
source 1..251
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
Db |||||||
175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 116
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||||
115 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 75
RESULT 7
HUMUT5345
LOCUS HUMUT5345 344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;
microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 344)
AUTHORS Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

```

```

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAAACAGGAGGCAAAATGC
Primer B: TTCGGGAAATGTCGCGGAC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2
FEATURES
source 1..344
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="8"
36..224
/standard_name="STS UT5345"
36..60
complement(202..224)
primer_bind
primer_bind
ORIGIN
Query Match 100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
Db |||||||
141 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 200
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||||
201 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 241
RESULT 8
BD195256/c
LOCUS BD195256 400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 400)
AUTHORS Dillon,P.J., Choi,G.H. and Welch,R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT
OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
PATRICK J DILLON,GIL H CHOI,RODNEY A WELCH
PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
  /organism='Unidentified'.
FEATURES
  source 1..400
    /organism='unidentified'
    /mol_type='genomic DNA'
    /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 165 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 106
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
Db 105 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
  beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
  synthetic construct
  other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
  Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
  WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
  EARTH CHEM CORP LTD
COMMENT OS Artificial gene
  OC Artificial sequence; Genes.
  PN JP 1986149089-A/1
  PD 07-JUL-1986
  PF 21-DEC-1984 JP 1984271206
  PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
  PI KITAZAWA NORIYUKI,
  PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
  C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
  C12R1:19);
  CC strandedness: Double;
  CC topology: Linear;
  CC hypothetical: No;
  CC anti-sense: No;
  CC *source: strain=HB101;
  CC *source: clone=pvG201;
  CC Feature is identified by experimental;
  FH Key Location/Qualifiers
  FH promoter 125..170
  FT of beta-lactamase
  FT RBS 200..203
  FT CDS 209..438
  FT /product='beta-urogastrone precursor' FT
  sig_peptide 209..277
  FT /product='signal peptide of beta-lactonase' FT
  mat_peptide 278..435
  FT /product='beta-urogastrone mature peptide'.
  FT Location/Qualifiers
  FT 1..456
  /organism='synthetic construct'
  /mol_type='genomic DNA'
FEATURES
  source

```

```

ORIGIN
  /db_xref='taxon:32630'
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
  synthetic construct
  other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
  BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
  EARTH CHEM CORP LTD
COMMENT OS Artificial gene
  OC Artificial sequence; Genes.
  PN JP 1987083890-A/1
  PD 17-APR-1987
  PF 09-OCT-1985 JP 1985225393
  PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
  KOIDE TAKAO,
  PI OKAI HIDEO
  PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
  C12R1:125);
  CC strandedness: Double;
  CC topology: Linear;
  CC hypothetical: No;
  CC anti-sense: No;
  CC *source: clone=pvG201;
  FH Key Location/Qualifiers
  FT promoter 125..170
  FT /note='beta lactamase promoter' FT RBS
  CDS 200..204
  FT /product='beta urogastrone'
  FT sig_peptide 209..277
  FT mat_peptide 278..436
  FT /product='beta urogastrone'.
  FT Location/Qualifiers
  FT 1..456
  /organism='synthetic construct'
  /mol_type='genomic DNA'
  /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101

```

Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
|||||  
E01274 456 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and  
signal peptide of beta-lactamase.  
ACCESSION E01274  
VERSION E01274.1 GI:2169533  
KEYWORDS JP 1987179398-A/1.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Okai, H., Kumakura, T., Kawamoto, S., Adachi, S., Matsubara, A.,  
Ojida, K., Yano, M., Mihara, S., Matsuhiro, A. and Yanai, N.  
TITLE PRODUCTION OF BETA-UROGASTRONE  
JOURNAL EARTH CHEM CORP LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
OS Homo sapiens  
PN JP 1987179398-A/1  
PD 06-AUG-1987  
PF 31-JAN-1986 JP 1986021032  
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,  
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,  
MATSUSHIRO AIZO, YANAIHARA NOBORU  
PC C12P21/00, C12N15/00, (C12P21/00, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No; Location/Qualifiers  
FH Key  
FH  
FT Promoter 125..170  
FT RBS 200..203  
FT sig\_peptide 209..277  
FT mat\_peptide 278..436  
FT CDS 209..439  
FT /product='beta-urogastron'.  
FT /product='beta-urogastron'.  
FEATURES  
source  
1. 456  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 456;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 60  
|||||  
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 114  
|||||  
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
|||||  
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
|||||  
RESULT 13  
AX260098/c  
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 60 from Patent WO0172774.  
ACCESSION AX260098  
VERSION AX260098.1 GI:16509129  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Drosophila melanogaster  
REFERENCE 1  
AUTHORS Deak, P., Glover, D.M. and Midgley, C.  
TITLE Cell cycle progression proteins  
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;  
Cyclacel Limited (GB)  
FEATURES  
source  
1. 466  
/organism='Drosophila melanogaster'  
/mol\_type='unassigned DNA'

Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
|||||  
E01274 456 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and  
signal peptide of beta-lactamase.  
ACCESSION E01274  
VERSION E01274.1 GI:2169533  
KEYWORDS JP 1987179398-A/1.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Okai, H., Kumakura, T., Kawamoto, S., Adachi, S., Matsubara, A.,  
Ojida, K., Yano, M., Mihara, S., Matsuhiro, A. and Yanai, N.  
TITLE PRODUCTION OF BETA-UROGASTRONE  
JOURNAL EARTH CHEM CORP LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
OS Homo sapiens  
PN JP 1987179398-A/1  
PD 06-AUG-1987  
PF 31-JAN-1986 JP 1986021032  
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,  
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,  
MATSUSHIRO AIZO, YANAIHARA NOBORU  
PC C12P21/00, C12N15/00, (C12P21/00, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No; Location/Qualifiers  
FH Key  
FH  
FT Promoter 125..170  
FT RBS 200..203  
FT sig\_peptide 209..277  
FT mat\_peptide 278..436  
FT CDS 209..439  
FT /product='beta-urogastron'.  
FT /product='beta-urogastron'.  
FEATURES  
source  
1. 456  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'  
ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 456;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 60  
|||||  
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 114  
|||||  
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
|||||  
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73  
|||||  
RESULT 12  
E01302/c  
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding  
promoter and signal peptide of beta-lactamase.  
ACCESSION E01302  
VERSION E01302.1 GI:2169561  
KEYWORDS JP 1987190083-A/1.  
SOURCE synthetic construct

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 221
    |||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||
Db 220 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 180

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION        AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1
AUTHORS
Deak, P., Glover, D.M. and Midgley, C.
TITLE
Cell cycle progression proteins
JOURNAL
Patent: WO 0172774-A 112 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..573
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296
    |||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||
Db 295 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 255

RESULT 15
A43586
LOCUS
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION        A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM
Cuphea lanceolata
Cuphea lanceolata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Myrtales; Lythraceae; Cuphea.
1 (bases 1 to 693)
Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
Schulte, W., Voetz, M., Walek, J. and Schell, J.
PROMOTERS
TITLE
Patent: WO 9507357-A 11 16-MAR-1995;
JOURNAL
MAX PLANCK GESELLSCHAFT (DB)
COMMENT
Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN.

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 592 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 651
    |||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||
Db 652 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 692
    |||||

Search completed: July 14, 2005, 14:03:36
Job time : 756.618 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_7131\_7231

Perfect score: 101

Sequence: 1 aggttattgtctcatgacg.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002as.\*

7: Geneseqn2002bs.\*

8: Geneseqn2003as.\*

9: Geneseqn2003bs.\*

10: Geneseqn2003cs.\*

11: Geneseqn2003ds.\*

12: Geneseqn2004as.\*

13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2 AAV76919	AAV76919 Staphyloc
C 2	101	100.0	228	1 AAN10032	Aan10032 Sequence
C 3	101	100.0	251	1 AAN10031	Aan10031 Sequence
C 4	101	100.0	400	2 AAV31229	AAV31229 E. coli J
C 5	101	100.0	456	1 AAN60624	Aan60624 Plasmid p
C 6	101	100.0	456	1 AAN71080	Aan71080 Sequence
C 7	101	100.0	456	1 AAN70833	Aan70833 Beta-urog
C 8	101	100.0	456	1 AAN81765	Aan81765 Sequence
C 9	101	100.0	466	6 ABA90413	ABa90413 Drosophil
C 10	101	100.0	487	2 AAX21173	Aax21173 Polynucle
C 11	101	100.0	535	2 AAX21149	Aax21149 Polynucle
C 12	101	100.0	573	6 ABA90456	ABa90456 Drosophil
C 13	101	100.0	605	12 ADH58311	ADH58311 Electroph
C 14	101	100.0	776	4 AAS30560	Aas30560 DNA encod
C 15	101	100.0	776	4 AAS27819	Aas27819 DNA encod
C 16	101	100.0	776	4 ABK42984	ABk42984 Genomic s
C 17	101	100.0	776	4 AAL07344	Aal07344 Human rep
C 18	101	100.0	776	4 AAL03229	Aal03229 Human rep
C 19	101	100.0	776	4 AAL06588	Aal06588 Human rep
C 20	101	100.0	776	4 AAL07340	Aal07340 Human rep

C 21	101	100.0	776	5 ABA14573	ABa14573 Human ner
C 22	101	100.0	776	5 AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8 ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8 ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8 ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9 ADB91869	ADB91869 Human sec
C 27	101	100.0	776	9 ADB61140	ADB61140 Connectiv
C 28	101	100.0	776	10 ADB94622	ADB94622 Novel hum
C 29	101	100.0	776	10 ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10 ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12 ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4 AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4 AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4 ABK42983	ABk42983 Genomic s
C 35	101	100.0	845	4 AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4 AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4 AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4 AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4 AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4 AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4 AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4 AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5 ABA14572	ABa14572 Human ner
C 44	101	100.0	845	5 AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9 ADB61139	ADB61139 Connectiv

#### ALIGNMENTS

##### RESULT 1

AAV76919/c

ID AAV76919 standard; DNA; 142 BP.

XX AC AAV76919;

DT 16-MAR-1999 (first entry)

XX DE Staphylococcus aureus contig SEQ ID #2608.

KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.

XX OS Staphylococcus aureus.

XX PN EP786519-A2.

XX PD 30-JUL-1997.

XX PF 07-JAN-1997; 97EP-00100117.

XX PR 05-JAN-1996; 96US-0009861P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX XX WPI; 1997-374922/35.

XX DR Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
XX stored on computer readable medium and used in the production of anti-  
XX S.aureus vaccines.

XX PS Claim 1; Page 2287; 3271pp; English.

XX CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
XX of the invention. The DNA sequences are recorded on a computer readable  
XX medium, preferably selected from a floppy or hard disk, random access  
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
XX the S.aureus DNA sequences allows putative functions to be assigned so  
XX that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can  
 CC be used in a vaccine composition against *S. aureus* infection. The  
 CC polypeptides can also be used in a kit for the immunodetection of  
 CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,  
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
 CC for recombinant production of the polypeptides. The new DNA sequences  
 CC (and their fragments) are useful as primers or probes for isolating  
 CC homologues of any of the *S. aureus* DNA sequences contained on the computer  
 CC readable medium

SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60  
 |||||  
 Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 48  
 |||||

Qy 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101  
 |||||  
 Db 47 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 7  
 |||||

RESULT 2  
 AAN10032/C  
 ID AAN10032 standard; DNA; 228 BP.  
 XX AC AAN10032;  
 XX 13-AUG-1992 (first entry)  
 DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.  
 XX Cloning vehicle; bacterial vector; transformed host; penicillinase;  
 KW insulin; ds.  
 XX Escherichia coli.  
 XX Key Location/Qualifiers  
 FT misc\_feature 1..4  
 FT /\*tag= a  
 FT /label= sticky end  
 FT misc\_feature 225..228  
 FT /\*tag= b  
 FT /label= sticky end  
 XX EP38182-A.  
 XX 21-OCT-1981.  
 XX 09-APR-1981; 81EP-00301561.  
 XX 11-APR-1980; 80US-00139225.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Gilbert W, Talmadge K;  
 XX WPI; 1981-80125D/44.  
 XX P-PSDB; AAP10039.  
 XX Synthesis of mature protein or polypeptide - by using bacterial host  
 XX transformed by cloned vehicle contg. DNA fragment etc.  
 XX Example; Fig 3; 34pp; English.  
 XX The closest identifiable promoter for the penicillinase gene in pKT241  
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was  
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the  
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20  
 CC nucleotides before its translational start signal. In the examples, the  
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
 CC fragment (CB6) for rat preproinsulin (see AAN10034)

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60  
 |||||  
 Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 116  
 |||||

Qy 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101  
 |||||  
 Db 115 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 75  
 |||||

RESULT 3  
 AAN10031/C  
 ID AAN10031 standard; DNA; 251 BP.  
 XX AC AAN10031;  
 XX 13-AUG-1992 (first entry)  
 DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.  
 XX Cloning vehicle; bacterial vector; transformed host; penicillinase;  
 KW insulin; ds.  
 XX Escherichia coli.  
 XX Key Location/Qualifiers  
 FT misc\_feature 1..4  
 FT /\*tag= a  
 FT /label= sticky end  
 FT misc\_feature 248..251  
 FT /\*tag= b  
 FT /label= sticky end  
 XX EP38182-A.  
 XX 21-OCT-1981.  
 XX 09-APR-1981; 81EP-00301561.  
 XX 11-APR-1980; 80US-00139225.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Gilbert W, Talmadge K;  
 XX WPI; 1981-80125D/44.  
 XX P-PSDB; AAP10038.  
 XX Synthesis of mature protein or polypeptide - by using bacterial host  
 XX transformed by cloned vehicle contg. DNA fragment etc.  
 XX Example; Fig 2; 34pp; English.  
 XX The closest identifiable promoter for the penicillinase gene in pKT241  
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its  
 CC translational start signal. In the examples, the 3' end of pKT241 was  
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the  
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20  
 CC nucleotides before its translational start signal. In the examples, the



CC 3' end of pK7218 was attached to the signal DNA sequence of the DNA  
 CC fragment (CB6) for rat preproinsulin (see AAN10034)  
 XX  
 SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 60  
 Db 175 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 116

Qy 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101  
 Db 115 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75

RESULT 4  
 AAV31229/c  
 ID AAV31229 standard; DNA; 400 BP.  
 AC AAV31229;  
 XX  
 DT 01-OCT-1998 (first entry)  
 XX  
 KW E. coli J96 pathogenicity island contig #43.  
 KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pher;  
 KW PAI V; pheV; vaccine; protective immune response; ds.  
 OS Escherichia coli.  
 XX  
 PN WO9822575-A2.  
 XX  
 PD 28-MAY-1998.  
 XX  
 PF 21-NOV-1997; 97WO-US021347.  
 XX  
 PR 22-NOV-1996; 96US-0031626P.  
 PR 14-OCT-1997; 97US-0061953P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (UWVI-) UNIV WISCONSIN.  
 XX  
 PI Dillon PJ, Choi GH, Welch RA;  
 XX  
 DR WPI; 1998-312461/27.  
 XX  
 PT New isolated uropathogenic E. coli nucleotide sequences - used to develop  
 PT products for the detection of pathogenic E. coli and to elicit an immune  
 PT response to pathogenic E. coli.  
 XX  
 PS Claim 21; Page 140-141; 250pp; English.  
 XX

This sequence represents a E. coli strain J96 contig containing  
 pathogenicity island (PAI) sequences, and represents a nucleic acid  
 molecule of the invention. PAIs are large fragments of DNA which comprise  
 pathogenicity determinants. The sequences of the invention are taken from  
 PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)  
 on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
 approximately 94 min (at pher) on the E. coli chromosome and is  
 approximately 160 kb in size. Antibodies specific to the proteins encoded  
 by the PAI open reading frames of the invention can be used in kits to  
 detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
 a protective immune response in an animal to the uropathogenic E. coli  
 strain J96

XX  
 SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 60  
 Db 165 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 106

Qy 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101  
 Db 105 GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC 65

RESULT 5  
 AAN60624/c  
 ID AAN60624 standard; DNA; 456 BP.  
 AC AAN60624;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 29-OCT-1991 (first entry)  
 XX  
 KW Plasmid pUG201 sequence encoding beta-urogastrone.  
 KW Beta-lactamase signal peptide; pGH54; pGH55; ss.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 125..170  
 FT promoter /\*tag= a  
 FT RBS 200..203  
 FT CDS /\*tag= b  
 209..439  
 FT sig\_peptide /\*tag= c  
 209..277  
 FT /\*tag= d  
 FT /label= Beta-lactamase signal peptide  
 278..436  
 FT mat\_peptide /\*tag= e  
 FT /label= Beta-urogastrone

W08603779-A.  
 XX  
 PN 03-JUL-1986.  
 XX  
 PD 19-DEC-1985; 85WO-JP000696.  
 XX  
 PR 21-DEC-1984; 84JP-00271206.  
 XX  
 PA (EART ) EARTH CHEM CO LTD.  
 PA (OHGA/) OHGAI H.  
 XX  
 PI Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;  
 XX  
 DR WPI; 1986-182911/28.  
 DR P-PSDB; AAP60678.  
 XX

Recombinant vector for polypeptide secretion - contains signal peptide  
 PT sequence directly bonded to peptide-coding sequence.  
 XX  
 PS Disclosure; Table 4; 79pp; Japanese.  
 XX

The plasmid produces secreted beta-urogastrone in a transformed  
 CC expression system. Similar plasmids may be constructed where the  
 CC secretion signal may be coupled with eg. somatostatin, insulin, growth  
 CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,  
 CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX

SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73

RESULT 6
AA71080/c
ID AA71080 standard; DNA; 456 BP.
XX
AC AA71080;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
KW pUGT 150s; beta-UG; ds.
XX
OS Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 125..170
FT promoter /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
PN JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
XX host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
XX The peptide may be expressed from plasmid pUGT 150s in a transformed
XX E.coli host. The plasmid may carry several separately expressing
XX sequences comprising a tac promoter, SD site, signal peptide, and coding
XX sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
XX add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73

RESULT 7
AA71080/c
ID AA71080 standard; DNA; 456 BP.
XX
AC AA71080;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
KW pUGT 150s; beta-UG; ds.
XX
OS Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 125..170
FT promoter /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
PN JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
XX host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
XX The peptide may be expressed from plasmid pUGT 150s in a transformed
XX E.coli host. The plasmid may carry several separately expressing
XX sequences comprising a tac promoter, SD site, signal peptide, and coding
XX sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
XX add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73

RESULT 8
AA71080/c
ID AA71080 standard; DNA; 456 BP.
XX
AC AA71080;
XX
DT 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

```

AA70833/c
ID AA70833 standard; DNA; 456 BP.
XX
AC AA70833;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
DE Beta-urogastrone sequence.
XX
KW Tumour; inosine; DNA probe; ds.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH 125..170
FT promoter /*tag= b
FT RBS 200..204
FT /*tag= c
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
PN JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
XX P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
XX using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
XX polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
XX ssDNA and probe are hybridized and the existence of DNA in the product is
XX detected. It can be used to detect the presence of malignant tumour.
XX (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
XX to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73

RESULT 8
AA71080/c
ID AA71080 standard; DNA; 456 BP.
XX
AC AA71080;
XX
DT 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

XX DE Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),  
XX Arg (53).  
XX KW Gastric acid secretion; cell proliferation; hormone; ds.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX CDS 209..277 /\*tag= a  
XX FT 278..439 /\*tag= b  
XX FT /\*tag= b  
XX FT /product= "New beta-urogastrone deriv."  
XX PN JP63012298-A.  
XX PD 19-JAN-1988.  
XX PF 30-JUN-1986; 86JP-00153783.  
XX PR 30-JUN-1986; 86JP-00153783.  
XX PA (EART ) EARTH SEIYAKU KK.  
XX DR WPI; 1988-054638/08.  
XX DR P-PSDB; AAP81349.  
XX PT New beta-urogastrone deriv. - has gastric acid secretion inhibition and  
XX proliferation promotion activity.  
XX PS Disclosure; Page 685; 76pp; Japanese.  
XX CC The deriv. has various biological activities such as gastric acid  
XX secretion inhibiting action, or cell proliferation promoting action. The  
XX deriv. has the same biological or pharmacological activities as beta-  
XX urogastrone. It is not susceptible to denaturation by oxidn. and is  
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as  
XX pepsinase. (Updated on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 60  
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 114  
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101  
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73  
RESULT 9  
ABA90413/c  
ID ABA90413 standard; DNA; 466 BP.  
XX AC ABA90413;  
XX DT 12-FEB-2002 (first entry)  
XX DE Drosophila cell cycle progression protein coding sequence #48.  
XX KW Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;  
XX antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;  
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;  
XX cell cycle progression protein; tumour; proliferative disorder;  
XX cardiovascular; autoimmune; dermatological disorder; ds.  
XX OS Drosophila sp.  
XX PI Fraser CM;  
XX

PN WO200172774-A2.  
XX PD 04-OCT-2001.  
XX PF 23-MAR-2001; 2001WO-GB001297.  
XX PR 24-MAR-2000; 2000GB-00007268.  
XX PA (CYCL-) CYCLACEL LTD.  
XX PI Deak P, Glover DM, Midgley C;  
XX DR WPI; 2002-055132/07.  
XX PT Polynucleotides encoding cell cycle progression proteins, useful for  
XX treating a tumor or a proliferative disorder.  
XX PS Claim 1; Page 99; 213pp; English.  
XX CC The present invention relates to Drosophila cell cycle progression  
XX proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-  
XX ABA90520). The coding sequences and proteins are useful for identifying a  
XX substance capable of affecting the function of the corresponding gene,  
XX inhibiting mitosis and/or meiosis. They can also be used in a method for  
XX treating a tumour or proliferative disorder, cardiovascular disorders  
XX (such as restenosis and cardiomyopathy), autoimmune disorders such as  
XX (glomerulonephritis and rheumatoid arthritis), dermatological disorders  
XX (such as psoriasis), antiinflammatory, antifungal and antiparasitic  
XX disorders (such as malaria)  
XX SQ Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;  
Query Match 100.0%; Score 101; DB 6; Length 466;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 60  
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 221  
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101  
Db 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180  
RESULT 10  
AAX21173/c  
ID AAX21173 standard; DNA; 487 BP.  
XX AC AAX21173;  
XX DT 05-MAY-1999 (first entry)  
XX DE Polynucleotide sequence from the genome of Treponema pallidum.  
XX KW Treponema pallidum infection; syphilis; Borrelia infection; animal;  
XX enzyme production; ds.  
XX OS Treponema pallidum.  
XX PN WO9859034-A2.  
XX PD 30-DEC-1998.  
XX PF 23-JUN-1998; 98WO-US013041.  
XX PR 24-JUN-1997; 97US-0050667P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Fraser CM;  
XX

DR WPI; 1999-081273/07.

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

XX Claim 1; Page 1106; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.6e-21; Length 487;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

Db 323 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 264

Qy 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101

Db 263 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 223

RESULT 11

AAX21149/c

ID AAX21149 standard; DNA; 535 BP.

XX

AC AAX21149;

XX

DT 05-MAY-1999 (first entry)

XX

XX Polynucleotide sequence from the genome of *Treponema pallidum*.

DE

XX *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;

KW enzyme production; ds.

XX

XX *Treponema pallidum*.

OS

XX WO9859034-A2.

XX

XX 30-DEC-1998.

XX

XX 23-JUN-1998; 98WO-US013041.

XX

XX 24-JUN-1997; 97US-0050667P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Fraser CM;

XX

XX WPI; 1999-081273/07.

XX

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

XX Claim 1; Page 1093; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;

Query Match 100.0%; Score 101; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.7e-21; Length 535;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296

Query Match 100.0%; Score 101; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.7e-21; Length 535;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296

RESULT 12

ABA90456/c

ID ABA90456 standard; DNA; 573 BP.

XX

AC ABA90456;

XX

DT 12-FEB-2002 (first entry)

XX

XX *Drosophila* cell cycle progression protein coding sequence #91.

DE

XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;

KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;

KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;

KW cell cycle progression protein; tumour; proliferative disorder;

KW cardiovascular; autoimmune; dermatological disorder; ds.

XX

XX *Drosophila* sp.

XX

XX WO200172774-A2.

XX

XX 04-OCT-2001.

XX

XX 23-MAR-2001; 2001WO-GB001297.

XX

XX 24-MAR-2000; 2000GB-00007268.

XX

XX (CYCL-) CYCLACEL LTD.

XX

XX Deak P, Glover DM, Midgley C;

XX

XX WPI; 2002-055132/07.

XX

XX Polynucleotides encoding cell cycle progression proteins, useful for

PT treating a tumor or a proliferative disorder.

XX

XX Claim 1; Page 144; 213pp; English.

XX

CC The present invention relates to *Drosophila* cell cycle progression

CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-

CC ABA90520). The coding sequences and proteins are useful for identifying a

CC substance capable of affecting the function of the corresponding gene, a

CC substance capable of inhibiting the cell division cycle, or capable of

CC inhibiting mitosis and/or meiosis. They can also be used in a method for

CC treating a tumour or proliferative disorder, cardiovascular disorders

CC (such as restenosis and cardiomyopathy), autoimmune disorders such as

CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders

CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic

CC disorders (such as malaria)

XX

SQ Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21; Length 573;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296



PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249247P.  
PR 17-NOV-2000; 2000US-0249248P.  
PR 17-NOV-2000; 2000US-0249249P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-476223/51.

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.  
Claim 1; SEQ ID NO 418; 512pp; English.

The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders of the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatodystonia, prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic hypertrophy or hyperplasia, and prostate neoplastic disorders, including adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (II) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders, urinary system disorders (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, disorders of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is useful to prevent skin aging, for preventing hair loss, to maintain organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;  
Best Local Similarity 100.0%; Pred. No. 2,9e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAATAG 60  
|||  
Db 546 AGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAATAG 487  
|||  
Qy 61 GGGTTCGGCGCACATTTCCCGAAAGTGCCACTGACGTC 101  
|||  
Db 486 GGGTTCGGCGCACATTTCCCGAAAGTGCCACTGACGTC 446  
|||

RESULT 15  
AAS27819/c  
ID AAS27819 standard; DNA; 776 BP.



PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465460/50.	
XX		
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or	
PT	prognosing disorders related to the proteins, including cancers, immune	
PT	disorders and neuronal disorders.	
XX		
XX	Claim 1; SEQ ID NO 1479; 880pp; English.	
XX		
CC	The invention relates to novel isolated polypeptides (I), and	
CC	polynucleotides (II). (I), (II) and the antibody to (I) are useful for	
CC	diagnosing, preventing and treating diseases including immune system	
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune	
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ	
CC	transplant rejections and graft versus host disease, infectious diseases	
CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and	
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative	
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders	
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities	
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.	
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),	
CC	respiratory disorders, dermatological disorders (e.g. wound healing,	
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's	
CC	disease), reproductive system disorders, gastrointestinal disorder	
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of	
CC	B-cell responsiveness to pathogens, activators of T-cells, to induce	
CC	higher affinity antibodies, and as a means to induce tumour proliferation	
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-	
CC	AAS27850 represent novel signal transduction pathway protein coding	
CC	sequences and PCR primers of the invention	
XX		
SQ	Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;	
	Query Match 100.0%; Score 101; DB 4; Length 776;	
	Best Local Similarity 100.0%; Pred. No. 2.9e-21;	
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGGGTTATTGTCCTAGTACGGCATACATATTGTAATGCTATTAGAAAAATAACAAATAG 60	
Db	546 AGGGTTATTGTCCTAGTACGGCATACATATTGTAATGCTATTAGAAAAATAACAAATAG 487	
QY	61 GGGTTCCCGCGCACATTTCCCCGAAAGTGGCCACTGACGTC 101	
Db	486 GGGTTCCCGCGCACATTTCCCCGAAAGTGGCCACTGACGTC 446	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)

3997.736 Million cell updates/sec

Title: US-09-482-682-64\_COPY\_7131\_7231

Perfect score: 101

Sequence: 1 aggggtattgtctcatgagc.....gaaagtgcacatgacgtc 101

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_ges1:\*

9: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CF766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

## ALIGNMENTS

RESULT 1  
BM078095/c  
LOCUS BM078095 300 bp mRNA linear EST 30-NOV-2001  
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma cylindrosporum cDNA 5', mRNA sequence.  
ACCESSION BM078095  
VERSION BM078095.1 GI:17157967  
KEYWORDS EST  
SOURCE Hebeloma cylindrosporum  
ORGANISM Hebeloma cylindrosporum  
Rukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Cortinariaceae; Hebeloma.  
REFERENCE 1 (bases 1 to 300)  
AUTHORS Wipf D., Benjdia M., Tegeder M. and Frommer W.B.  
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wipf D  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: pDR196 5' primer (PMA 5')  
High quality sequence stop: 300  
POLYA=No.

## FEATURES

source  
1..300  
Location/Qualifiers  
/organism="Hebeloma cylindrosporum"  
/mol\_type="mRNA"  
/strain="H1"  
/db\_xref="taxon:76867"  
/tissue\_type="Mycelia"  
/lab\_host="E. coli XLI-Blue"  
/clone\_lib="Hebeloma cylindrosporum functional cDNA library"  
/notes="Vector: pDR 196 (unpublished); Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 8.1e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGTTATTCTCATGCGGATACATATTGTAATGCTATTAGAAAAATAACAATAG 60

```

|||||
174 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 115
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1; ampicillinase (1e-10).
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        1..309
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="010H20aC4"
            /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 98
    |||
QY 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||
Db 99 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 139
    |||

RESULT 5
AL597149      391 bp      mRNA      linear      EST 04-SEP-2003
LOCUS
DEFINITION
DKFP313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
AL597149
ACCESSION
AL597149
VERSION
AL597149.1 GI:15154845
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 391)
AUTHORS
Koehler,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
TITLE
EST (Koehler,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL
Unpublished (1999)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BMPZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFP313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        1..391
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFP313J1611"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="313 (synonym: hlcc2)"
            /note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||
Db 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        1..309
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="010H20aC4"
            /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 98
    |||
QY 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||
Db 99 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 139
    |||

RESULT 6
CC819240/c
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 414)
AUTHORS
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL
Unpublished (2003)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 Row: D Column: 19
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        1..414
            /organism="Sterkiella histriomuscorum"
            /mol_type="genomic DNA"
            /db_xref="taxon:94289"
            /clone="UUGC100005D19"
            /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
            /clone_lib="Oxytricha plasmid UUGC10 library"
            /note="Vector: FWD42nv; Purified macronuclear chromosomal
            DNA from Oxytricha trifallax was blunt end-repaired with
            T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
            oligonucleotides were ligated to the blunt ends in high
            molar excess. Vector DNA was prepared from a derivative of
            pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
            derivative of plasmid R1. The vector was ligated with
            adaptors complementary to the insert adaptors and
            purified. The sheared, adaptor mouse DNA was annealed to
            adaptor vector DNA, and transformed into
            chemically-competent E. Coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||
Db 414 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 355
    |||

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        1..309
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="010H20aC4"
            /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||
Db 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAATAG 60
    |||

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCBST library Haplochromis chilotes cDNA clone no90C12,
mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Labroidae; Cichlidae; Haplochromis.
1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
Orig sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
source
1..417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stage="varied"
/clone_lib="HCBST library"

ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 60
|||||
DB 129 AGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 70
|||||

QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
DB 69 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 29
|||||

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10006J13 R, genomic survey
sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE
ORGANISM Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 Row: J Column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES
Location/Qualifiers
source
1..491
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10006J13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/vector="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 60
|||||
DB 412 AGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 353
|||||

QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
DB 352 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 312
|||||

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 495 bp mRNA linear EST 02-OCT-2001
S035A01 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES
Location/Qualifiers
source
1..495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

```

```
/tissue_type="Stem"
/dev_stages="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 60
    |||||||
Db 62 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 121

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 122 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 162

RESULT 10
CC818374/c
LOCUS
DEFINITION
CC818374 495 bp DNA linear GSS 17-JUL-2003
100004B07R Oxytricha plasmid UUGC100004B07 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 495)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
FEATURES
    source
    Location/Qualifiers
        1..495
            /organism="Sterkiella histriomuscorum"
            /mol_type="genomic DNA"
            /db_xref="taxon:94289"
            /clone="UUGC100004B07"
            /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
            /clone_lib="Oxytricha plasmid UUGC10 library"
            /note="Vector: PWD42nv; Purified macronuclear chromosomal
            DNA from Oxytricha trifallax was blunt end-repaired with
            T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
            oligonucleotides were ligated to the blunt ends in high
            molar excess. Vector DNA was prepared from a derivative of
            PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible
            derivative of plasmid R1. The vector was ligated with
            adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adaptor complementary to the insert adaptors and
            chemically-competent E. Coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 60
    |||||||
Db 391 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
```

```

Db      331 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 291
|||||
CC819854          503 bp      DNA      linear      GSS 17-JUL-2003
CC819854/c       100006N08R Oxytricha plasmid UUGC10 library Sterkiella
LOCUS           histriomuscorum genomic clone UUGC100006N08 R, genomic survey
DEFINITION      sequence.
ACCESSION      CC819854.1 GI:32900533
VERSION        CC819854
KEYWORDS       GSS.
SOURCE         Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM       Sterkiella histriomuscorum
               Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
               Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 503)
AUTHORS        Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE          Paired end reads from plasmid inserts of Oxytricha trifallax
               macronuclear chromosomes
JOURNAL        Unpublished (2003)
COMMENT        Contact: Robert B. Weiss
               University of Utah Genome Center
               University of Utah
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Plate: 0006 row: N column: 08
               Seq primer: CACACAGGAACACGTATGACC
               Class: plasmid ends
               High quality sequence stop: 503.
FEATURES       source
               Location/Qualifiers
               1..503
               /organism="Sterkiella histriomuscorum"
               /mol_type="genomic DNA"
               /db_xref="taxon:94289"
               /clones="UUGC100006N08"
               /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
               /note="Vector: PWD42nv; Purified macronuclear chromosomal
               DNA from Oxytricha trifallax was blunt end-repaired with
               T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
               oligonucleotides were ligated to the blunt ends in high
               molar excess. Vector DNA was prepared from a derivative of
               pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
               derivative of plasmid R1. The vector was ligated with
               adaptors complementary to the insert adaptors and
               purified. The sheared, adaptor mouse DNA was annealed to
               adaptor vector DNA, and transformed into
               chemically-competent E. Coli XL10-Gold (Stratagene) cells
               and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
|||||
Db      410 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 351
|||||
Qy      61 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 101
|||||
Db      350 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 310
|||||
RESULT 13
CC817752/c       100003C16R Oxytricha plasmid UUGC10 library Sterkiella
LOCUS           Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
DEFINITION      sequence.
ACCESSION      CC817752.1 GI:328996415
VERSION        CC817752
KEYWORDS       GSS.
SOURCE         Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM       Sterkiella histriomuscorum
               Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
               Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 515)
AUTHORS        Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE          Paired end reads from plasmid inserts of Oxytricha trifallax
               macronuclear chromosomes
JOURNAL        Unpublished (2003)
COMMENT        Contact: Robert B. Weiss
               University of Utah Genome Center
               University of Utah
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Plate: 0003 row: C column: 16
               Seq primer: CACACAGGAACACGTATGACC
               Class: plasmid ends
               High quality sequence stop: 515.
FEATURES       source
               Location/Qualifiers
               1..515
               /organism="Sterkiella histriomuscorum"
               /mol_type="genomic DNA"
               /db_xref="taxon:94289"
               /clones="UUGC100003C16"
               /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
               /note="Vector: PWD42nv; Purified macronuclear chromosomal
               DNA from Oxytricha trifallax was blunt end-repaired with
               T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
               oligonucleotides were ligated to the blunt ends in high
               molar excess. Vector DNA was prepared from a derivative of
               pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
               derivative of plasmid R1. The vector was ligated with
               adaptors complementary to the insert adaptors and
               purified. The sheared, adaptor mouse DNA was annealed to
               adaptor vector DNA, and transformed into
               chemically-competent E. Coli XL10-Gold (Stratagene) cells
               and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
|||||
Db      412 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 353
|||||
Qy      61 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 101
|||||
Db      352 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 312
|||||
RESULT 14
CC817128/c       100002D21R Oxytricha plasmid UUGC10 library Sterkiella
LOCUS           histriomuscorum genomic clone UUGC100002D21 R, genomic survey
DEFINITION      sequence.
ACCESSION      CC817128.1 GI:328996415
VERSION        CC817128
KEYWORDS       GSS.
SOURCE         Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM       Sterkiella histriomuscorum
               Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
               Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 515)
AUTHORS        Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE          Paired end reads from plasmid inserts of Oxytricha trifallax
               macronuclear chromosomes
JOURNAL        Unpublished (2003)
COMMENT        Contact: Robert B. Weiss
               University of Utah Genome Center
               University of Utah
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Plate: 0003 row: C column: 16
               Seq primer: CACACAGGAACACGTATGACC
               Class: plasmid ends
               High quality sequence stop: 515.
FEATURES       source
               Location/Qualifiers
               1..515
               /organism="Sterkiella histriomuscorum"
               /mol_type="genomic DNA"
               /db_xref="taxon:94289"
               /clones="UUGC100003C16"
               /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
               /note="Vector: PWD42nv; Purified macronuclear chromosomal
               DNA from Oxytricha trifallax was blunt end-repaired with
               T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
               oligonucleotides were ligated to the blunt ends in high
               molar excess. Vector DNA was prepared from a derivative of
               pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
               derivative of plasmid R1. The vector was ligated with
               adaptors complementary to the insert adaptors and
               purified. The sheared, adaptor mouse DNA was annealed to
               adaptor vector DNA, and transformed into
               chemically-competent E. Coli XL10-Gold (Stratagene) cells
               and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
|||||
Db      412 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 353
|||||
Qy      61 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 101
|||||
Db      352 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 312
|||||
RESULT 14
CC817128/c       100002D21R Oxytricha plasmid UUGC10 library Sterkiella
LOCUS           histriomuscorum genomic clone UUGC100002D21 R, genomic survey
DEFINITION      sequence.
ACCESSION      CC817128.1 GI:328996415
VERSION        CC817128
KEYWORDS       GSS.
SOURCE         Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM       Sterkiella histriomuscorum
               Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
               Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 515)
AUTHORS        Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE          Paired end reads from plasmid inserts of Oxytricha trifallax
               macronuclear chromosomes
JOURNAL        Unpublished (2003)
COMMENT        Contact: Robert B. Weiss
               University of Utah Genome Center
               University of Utah
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Plate: 0003 row: C column: 16
               Seq primer: CACACAGGAACACGTATGACC
               Class: plasmid ends
               High quality sequence stop: 515.
FEATURES       source
               Location/Qualifiers
               1..515
               /organism="Sterkiella histriomuscorum"
               /mol_type="genomic DNA"
               /db_xref="taxon:94289"
               /clones="UUGC100003C16"
               /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
               /note="Vector: PWD42nv; Purified macronuclear chromosomal
               DNA from Oxytricha trifallax was blunt end-repaired with
               T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
               oligonucleotides were ligated to the blunt ends in high
               molar excess. Vector DNA was prepared from a derivative of
               pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
               derivative of plasmid R1. The vector was ligated with
               adaptors complementary to the insert adaptors and
               purified. The sheared, adaptor mouse DNA was annealed to
               adaptor vector DNA, and transformed into
               chemically-competent E. Coli XL10-Gold (Stratagene) cells
               and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
|||||
Db      412 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 353
|||||
Qy      61 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 101
|||||
Db      352 GGGTTCGCGCACATTTCCTCCGAAAGTGCCACCTGACGTC 312
|||||

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Stichotrichida; Oxytrichidae; Sterkiella.  
1 (bases 1 to 518)  
Dunn,D., Doak,T., Herrick,G. and Weiss,R.  
Paired end reads from plasmid inserts of Oxytricha trifallax  
macronuclear chromosomes  
Unpublished (2003)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Plate: 0002 row: D column: 21  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 518.

FEATURES  
source  
1..518  
Location/Qualifiers  
/organism="Sterkiella histriomuscorum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:94289"  
/clone="UUGC100002J19"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Oxytricha plasmid UUGC10 library"  
/note="Vector: PWD42nv; Purified macronuclear chromosomal DNA from Oxytricha trifallax was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. Coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 101; DB 9; Length 518;  
Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAACAAATAG 60  
|||||  
Db 410 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAACAAATAG 351  
|||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101  
|||||  
Db 350 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 310  
|||||

RESULT 15  
CC817162/c  
LOCUS  
DEFINITION  
100002J19R Oxytricha plasmid UUGC10 library Sterkiella  
histriomuscorum genomic clone UUGC100002J19 R, genomic survey  
sequence.  
CC817162  
CC817162.1 GI:32896449  
GSS.  
Sterkiella histriomuscorum (Oxytricha trifallax)  
Sterkiella histriomuscorum  
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;  
Stichotrichida; Oxytrichidae; Sterkiella.  
1 (bases 1 to 519)  
Dunn,D., Doak,T., Herrick,G. and Weiss,R.  
Paired end reads from plasmid inserts of Oxytricha trifallax  
macronuclear chromosomes  
Unpublished (2003)  
Contact: Robert B. Weiss  
University of Utah Genome Center

University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Plate: 0002 row: J column: 19  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 519.

FEATURES  
source  
1..519  
Location/Qualifiers  
/organism="Sterkiella histriomuscorum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:94289"  
/clone="UUGC100002J19"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Oxytricha plasmid UUGC10 library"  
/note="Vector: PWD42nv; Purified macronuclear chromosomal DNA from Oxytricha trifallax was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. Coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 101; DB 9; Length 519;  
Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAACAAATAG 60  
|||||  
Db 416 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTAGAAAAATAACAAATAG 357  
|||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101  
|||||  
Db 356 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 316  
|||||

Search completed: July 14, 2005, 23:23:20  
Job time : 962.667 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds  
(without alignments)

6468.225 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_1\_100

Perfect score: 100

Sequence: 1 ctgtccctgtgtgtgtt.....caattgatgaagaatctgc 100

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	6	AX643583 Sequence
2	100	100.0	633	14	ALRPROLTB
3	100	100.0	648	6	AX175190 Sequence
4	100	100.0	648	6	AX175195 Sequence
5	100	100.0	1070	6	AB5308 Sequence 6
6	100	100.0	1070	6	BD107647 FIV vacci
7	100	100.0	2245	6	AX643582 Sequence
8	100	100.0	2426	6	AX044426 Sequence
9	100	100.0	2427	6	AX044425 Sequence
10	100	100.0	3557	12	SYNRSV3MV
11	100	100.0	3840	12	EV132038
12	100	100.0	3853	6	AR098190 Sequence
13	100	100.0	3853	6	AR207832 Sequence
14	100	100.0	3853	6	BD009729 Tissue sp
15	100	100.0	3925	6	A60213 Sequence 9
16	100	100.0	3925	6	ARI22289
17	100	100.0	3986	12	PCDNA32EO
18	100	100.0	4026	6	AR098191 Sequence
19	100	100.0	4026	6	AR207833 Sequence

20	100	100.0	4026	6	BD009730
21	100	100.0	4059	6	AR071324 Sequence
22	100	100.0	4249	6	AR098192 Sequence
23	100	100.0	4249	6	AR207834 Sequence
24	100	100.0	4249	6	BD009731 Tissue sp
25	100	100.0	4341	6	A38214 Sequence 58
26	100	100.0	4341	6	AX286570 Sequence
27	100	100.0	4457	6	AX743954 Sequence
28	100	100.0	4525	6	AR062871 Sequence
29	100	100.0	4597	6	AX060344 Sequence
30	100	100.0	4839	12	SYNRSV5GPT
31	100	100.0	4840	6	AX133940 Sequence
32	100	100.0	4965	6	AR071323 Sequence
33	100	100.0	5053	6	BD238492 Expressio
34	100	100.0	5070	6	AX234391 Sequence
35	100	100.0	5082	6	A91754 Sequence 10
36	100	100.0	5082	6	BD085110 Vertebrat
37	100	100.0	5108	12	SYNRSV5NEO
38	100	100.0	5162	6	AX951626 Sequence
39	100	100.0	5257	12	CVU89673
40	100	100.0	5432	6	BD234590 Screening
41	100	100.0	5432	6	AX026821 Sequence
42	100	100.0	5446	6	BD195386 Compositi
43	100	100.0	5446	6	AX319694 Sequence
44	100	100.0	5564	12	SYNTRC
45	100	100.0	5618	6	A44171 Sequence 1

#### ALIGNMENTS

RESULT 1  
LOCUS AX643583 562 bp DNA linear PAT 24-FEB-2003  
DEFINITION Sequence 2 from Patent WO02099100.  
ACCESSION AX643583  
VERSION AX643583.1 GI:28551383  
KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
REFERENCE 1  
AUTHORS Al-Rubeai, M. and Shuttleworth, J.  
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21  
JOURNAL Patent: WO 02099100-A 2 12-DEC-2002;  
Lonza Biologics plc (GB)  
FEATURES  
Location/Qualifiers  
source 1..562  
/organism="Mus sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10095"  
/note="Rous Sarcoma Virus LTR promoter"

ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 562;  
Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGTTGTGTGGAGGTGCTGAGTAGTGCAGCGACGACAAATTAAGCTACA 60  
Db 46 CTGCTCCCTGTTGTGTGGAGGTGCTGAGTAGTGCAGCGACGACAAATTAAGCTACA 105  
Qy 61 ACAAGGCAAGCTTGACCGCAATTGTCATGAAGAATCTGC 100.  
Db 106 ACAAGGCAAGCTTGACCGCAATTGTCATGAAGAATCTGC 145

RESULT 2  
ALRPROLTB 633 bp ss-RNA linear VRL 28-APR-1993  
LOCUS ALRPROLTB  
DEFINITION Rous sarcoma virus (Schmidt-Ruppin), proviral, 3' LTR on 21S mRNA.

ACCESSION J02025 J02022  
VERSION J02025.1 GI:210255  
KEYWORDS c-myc proto-oncogene; long terminal repeat (LTR); src oncogene.  
SOURCE Rous sarcoma virus  
ORGANISM Rous sarcoma virus  
REFERENCE 1 (sites)  
AUTHORS Yamamoto, T., de Crombrughe, B. and Pastan, I.  
TITLE Identification of a functional promoter in the long terminal repeat  
of Rous sarcoma virus  
JOURNAL Cell 22 (3), 787-797 (1980)  
MEDLINE 81112147  
PUBMED 6257399  
REFERENCE 2 (bases 1 to 633)  
AUTHORS Yamamoto, T., Tyagi, J.S., Fagan, J.B., Jay, G., deCrombrughe, B. and Pastan, I.  
TITLE Molecular mechanism for the capture and excision of the transforming gene of avian sarcoma virus as suggested by analysis of recombinant clones  
JOURNAL J. Virol. 35 (2), 436-443 (1980)  
MEDLINE 81072438  
PUBMED 6255184  
REFERENCE 3 (bases 319 to 633)  
AUTHORS Yamamoto, T., Jay, G. and Pastan, I.  
TITLE Unusual features in the nucleotide sequence of a cDNA clone derived from the common region of avian sarcoma virus messenger RNA  
Proc. Natl. Acad. Sci. U.S.A. 77 (1), 176-180 (1980)  
JOURNAL 80145590  
MEDLINE 6244542  
PUBMED  
COMMENT Original source text: Rous sarcoma virus (Schmidt-Ruppin strain, subgroup D) provirus, cDNA to 21S mRNA from infected chicken embryonic fibroblasts, clone pSR1.  
[1] sites; mRNA start.  
Original figure in [2] included 24 'g's on 5' end and 16 'c's on 3' end that were cDNA synthesis artifacts.  
[2] also sequenced a defective clone, pSR2, with the src gene deleted (see separate entry).  
[1] demonstrated the mRNA transcription initiation site shown in the Sites table using pSR1 as a template. However, this is the 3' LTR, and the functional mRNA start site would be assumed to be on the 5' LTR at the homologous site.

FEATURES  
source  
1..633  
/organism="Rous sarcoma virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:11886"  
misc\_RNA  
<1..517  
/note="viral genomic RNA"  
LTR  
211..5633  
/note="3' LTR"  
mRNA  
517..5633  
/note="in vitro mRNA [1]; see comment"  
repeat\_region  
517..536  
/note="terminally redundant repeat"  
ORIGIN  
20 bp upstream of PstI site.

Query Match 100.0%; Score 100; DB 14; Length 633;  
Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 60  
|||||  
Db 28 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 87  
|||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 88 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 127  
|||||

RESULT 3  
AX175190  
LOCUS  
DEFINITION Sequence 1 from Patent WO0142444.

AX175190  
VERSION AX175190.1 GI:14598581  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Rivera, V., Zoltick, P. and Wilson, J.M.  
TITLE Methods for expression of genes in primates  
JOURNAL Patent: WO 0142444-A 1 14-JUN-2001;  
ARIAD GENE THERAPEUTICS, INC. (US); THE UNIVERSITY OF PENNSYLVANIA (US)  
FEATURES  
source  
1..648  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="vector/RSV promoter/vector"  
ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 648;  
Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 60  
|||||  
Db 90 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 149  
|||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189  
|||||

RESULT 4  
AX175195  
LOCUS  
DEFINITION Sequence 6 from Patent WO0142444.  
ACCESSION AX175195  
VERSION AX175195.1 GI:14598586  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Rivera, V., Zoltick, P. and Wilson, J.M.  
TITLE Methods for expression of genes in primates  
JOURNAL Patent: WO 0142444-A 6 14-JUN-2001;  
ARIAD GENE THERAPEUTICS, INC. (US); THE UNIVERSITY OF PENNSYLVANIA (US)  
FEATURES  
source  
1..648  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="MLuI/RSV promoter/BglI"  
ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 648;  
Best Local Similarity 100.0%; Pred. No. 2.6e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 60  
|||||  
Db 90 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTCGCGAGCAAAATTTAAGCTACA 149  
|||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
|||||  
Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189  
|||||

RESULT 5  
AX175190  
LOCUS  
DEFINITION Sequence 1 from Patent WO0142444.

```

DEFINITION Sequence 6 from Patent WO9840493.
ACCESSION A85308
VERSION A85308.1 GI:6733916
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 60
Db |||||||
QY 77 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 136
Db |||||||

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db |||||||
137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

RESULT 7
LOCUS AX643582 2245 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 1 from Patent WO2099100.
ACCESSION AX643582
VERSION AX643582.1 GI:28551382
KEYWORDS
SOURCE
ORGANISM Mus sp.
REFERENCE
AUTHORS Al-Rubeai, M. and Shuttleworth, J.
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 1 12-DEC-2002; Lonza Biologics plc (GB)
FEATURES
    source
        Location/Qualifiers
            1..2245
                /organism="Mus sp."
                /mol_type="unassigned DNA"
                /db_xref="taxon:10095"
                /note="RSV-LTR promoter + intron + p21 cds + Tkpoly(A) LacSwitch II expression construct"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 2245;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 60
Db |||||||
46 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 105
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db |||||||
106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145

RESULT 8
LOCUS AX044426 2426 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 18 from Patent WO0066752.
ACCESSION AX044426
VERSION AX044426.1 GI:11343299
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 18 09-NOV-2000; Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
    source
        Location/Qualifiers
            1..2426
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="CPG2 with last exon of Thy-1 fused at 3' end"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 60
Db |||||||
77 CTGCTCCCTGCTGTGTGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTACA 136
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db |||||||
137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

RESULT 6
LOCUS BD107647 1070 bp DNA linear PAT 18-SEP-2002
DEFINITION FIV vaccine.
ACCESSION BD107647
VERSION BD107647.1 GI:23202465
KEYWORDS JP 2002501369-A/6.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
AUTHORS Neil, J.C., Rigby, M.A. and Jarrett, J.O.
TITLE FIV vaccine
JOURNAL Patent: JP 2002501369-A 6 15-JAN-2002; THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW
COMMENT OS Artificial Sequence
PN JP 2002501369-A/6
PD 15-JAN-2002
PF 10-MAR-1998 JP 1998539351
PI 11-MAR-1997 GB 9704977.9
PJ JAMES CHARLES NEIL, MARK ALAN RIGBY, JAMES OSWALD JARRETT PC
CI 2N15/49, A61K31/70, A61K48/00
CC CMV PROMOTER FROM pCDNA3 (a Bgl II - Kpn I restriction fragment)
CC SST I - SST I FRAGMENT IN PLASMID CMV DEL. RT CC FIV GENOME
FROM THE t-RNA PRIMER BINDING SITE TO THE VIRAL Sst IS
CC
FT Key Location/Qualifiers
FT source 1..1070
FT /organism="Artificial Sequence".
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
ORIGIN

```

```

Query Match      100.0%; Score 100; DB 6; Length 2426;
Best Local Similarity 100.0%; Pred. No. 3.1e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 60
Db 69 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 128

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
Db 129 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 168

RESULT 9
LOCUS AX044425 2427 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 17 from Patent WO0066752.
ACCESSION AX044425
VERSION AX044425.1 GI:11343298
KEYWORDS synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 17 09-NOV-2000;
Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
source
1.2427
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="CPG2 mutant with last exon of Thy-1 fused at 3' end"

ORIGIN

Query Match      100.0%; Score 100; DB 6; Length 2427;
Best Local Similarity 100.0%; Pred. No. 3.1e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 60
Db 70 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 129

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
Db 130 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 169

RESULT 10
LOCUS SYNRSV3MV 3557 bp DNA circular SYN 27-APR-1993
DEFINITION Cloning vector RSV3.
ACCESSION M83240
VERSION M83240.1 GI:209303
KEYWORDS cDNA expression vector.
SOURCE unidentified cloning vector
ORGANISM unidentified cloning vector
other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS Messing, J.
TITLE New M13 vectors for cloning
JOURNAL Meth. Enzymol. 101, 20-78 (1983)
MEDLINE 83296918
PUBMED 6310323
REFERENCE 2
AUTHORS Gorman, C., Padmanabhan, R. and Howard, B.H.
TITLE High efficiency DNA-mediated transformation of primate cells
JOURNAL Science 221 (4610), 551-553 (1983)
MEDLINE 83249156
PUBMED 6306768
REFERENCE 3
  
```

```

AUTHORS Jacobson, S., Sekaly, R.P., Jacobson, C.L., McFarland, H.F. and Long, F.O.
TITLE HLA class II-restricted presentation of cytoplasmic measles virus antigens to cytotoxic T cells
JOURNAL J. Virol. 63 (4), 1756-1762 (1989)
MEDLINE 89178863
PUBMED 2784508
COMMENT Original source text: Cloning vector DNA.
FEATURES
source
Location/Qualifiers
1.3557
/organism="unidentified cloning vector"
/mol_type="genomic DNA"
/db_xref="taxon:45196"
misc_feature
1.29
/function="polylinker"
/evidence=experimental
misc_feature
912..3029
/function="ampicillin-resistance, replication origin"
/evidence=experimental
enhancer
3030..3557
/standard_name="5' LTR of Rous Sarcoma Virus"
/citation=[2]
/evidence=experimental

ORIGIN

Query Match      100.0%; Score 100; DB 12; Length 3557;
Best Local Similarity 100.0%; Pred. No. 3.2e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 60
Db 3030 CTGCTCCCTGCTTGTGTGTTGGAGTCCCTCAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 3089

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
Db 3090 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 3129

RESULT 11
LOCUS EVE132038 3840 bp RNA circular SYN 28-JUL-1999
DEFINITION Expression vector pCDPT.
ACCESSION AJ132038
VERSION AJ132038.1 GI:5640088
KEYWORDS AMP gene; beta lactamase; Cole1 origin of replication; multiple cloning site; SP6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; Xanthine-guanine phosphoribosyl transferase gene.
SOURCE Expression vector pCDPT
ORGANISM Expression vector pCDPT
other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS Zeng, B.J.
TITLE Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase tag
JOURNAL Unpublished
REFERENCE 2
AUTHORS Zeng, B.J.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA
FEATURES
source
Location/Qualifiers
1..3840
/organism="Expression vector pCDPT"
/mol_type="other RNA"
/db_xref="taxon:90749"
promoter
209..863
/note="CMV"
promoter
864..882
/note="T7"
misc_feature
882..984
/note="Multiple cloning site"
  
```

```
CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSEKIVTVMDLQIHARKLASRLMPSEQWKGIIAVSRGGLVPGA
LLARELGHVDVTCISYDHNQRELKVLKRAEGDGEFVIDLDVDTGTAVAIRE
MYPKAFHTTIFAKPAGRLVDYVDIPQDTWIEQPMGVMVFVPPISGR"
1649..1863
/note="BGH"
2450..2775
/note="SP6"
2644..2729
/note="SV40"
complement(2844..3704)
/gene="amp"
complement(2844..3704)
/gene="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHFRVALIPFAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILESFRPEFPMSTFKVLLCGAVLSRIDAGEQLRIHYSQNDLVE
YSPVTEKHLTDGMTRELCSSAATWMSDNTAANLLLTIGPKELTAFLHNNADKVTPL
DRWPELNEATPNDERTTMPVAMATLRKLLTGELLTLASRQQLTLDMWEADKVAQPL
LRSALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSAQTWDERNRQIA
EIGASLXKHW"
3632..>3840
/note="ColE1"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 6 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 65

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140
```

```
Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180
```

```
RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"
```

```
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180
```

```
RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
```

```
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCTGCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140
```



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	100	100.0	562	8	ABZ23250	Nucleotid
2	100	100.0	648	4	AAH43951	Abz43951
3	100	100.0	1070	2	AAV58058	Plasmid C
4	100	100.0	1506	12	ADM41035	Fungus n
5	100	100.0	1600	2	ADH11349	Vertebra
6	100	100.0	1782	12	ADM41037	Cytomega
7	100	100.0	2241	12	ADM41034	Human nu
8	100	100.0	2245	8	ABZ23249	Lac repre
9	100	100.0	2294	12	ADM41036	Cytomega
10	100	100.0	2426	4	AAD02037	Plasmid p
11	100	100.0	2427	4	AAD02036	Plasmid p
12	100	100.0	3400	2	AAT62937	3F4 huma
13	100	100.0	3400	2	AAT62932	2A2 huma
14	100	100.0	3853	2	AAV40006	Plasmid p
15	100	100.0	3925	2	AAT90695	Plasmid C
16	100	100.0	4026	2	AAV40007	Plasmid p
17	100	100.0	4059	2	AAQ75974	phLA-B7 e
18	100	100.0	4249	2	AAV63466	Plasmid p
19	100	100.0	4341	2	AAQ62391	Vector pV
20	100	100.0	4341	6	AA817704	Vector pV

CC particularly useful for maximizing or enhancing the production of e.g.  
 CC therapeutic proteins at an industrial scale  
 SQ Sequence 562 BP; 143 A; 109 C; 163 G; 147 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 100; DB 8; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-28;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 60  
 DB 46 CTGCTCCCTGCTTGTGTGTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 105  
 QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
 DB 106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145  
 RESULT 3  
 AAH43951  
 ID AAH43951 standard; DNA; 648 BP.  
 AC AAH43951;  
 DT 06-SEP-2001 (first entry)  
 DE Rous sarcoma virus promoter nucleotide sequence SEQ ID NO:1.  
 KW Rous sarcoma virus; promoter; enhancer; RSV; primate; gene expression;  
 KW transgene; genetic engineering; gene therapy; immunisation; ds.  
 XX Rous sarcoma virus.  
 XX WO200142444-A2.  
 XX 14-JUN-2001.  
 XX 08-DEC-2000; 2000WO-US033256.  
 XX 10-DEC-1999; 99US-0170019P.  
 XX (ARIA-) ARIAD GENE THERAPEUTICS INC.  
 XX (UYPE-) UNIV PENNSYLVANIA.  
 XX Rivera V, Zoltick P, Wilson JM;  
 XX WPI; 2001-381673/40.  
 XX Genetically engineering a primate for expression of a desired gene,  
 PT comprises introducing into the primate a transgene comprising Rous  
 PT Sarcoma Virus (RSV) promoter and a nucleic acid sequence heterologous to  
 PT RSV promoter.  
 XX Claim 7; Page 44; 64pp; English.  
 CC The present invention describes a method for genetically engineering a  
 CC primate for expression of a desired gene comprising introducing into the  
 CC primate a transgene comprising an Rous Sarcoma Virus (RSV) promoter and a  
 CC nucleic acid sequence heterologous to RSV promoter. Also described is a  
 CC primate cell (I) containing and capable of expressing a transgene  
 CC comprising an RSV promoter operably linked to a recombinant nucleic acid  
 CC encoding one or more fusion proteins, where the fusion proteins bind to a  
 CC ligand and in the presence of the ligand modulate(s) the expression level  
 CC of a target gene. The method can be used for high level expression of  
 CC genes in primates or for engineering primate cells. It is useful for  
 CC increasing the efficacy of many gene therapy strategies, and for  
 CC like ribozymes, antisense RNA, and dominant negative proteins, that act  
 CC either stoichiometrically, or by competition. The method increases the  
 CC efficacy of many gene therapy strategies by substantially elevating the  
 CC expression of an exogenous therapeutic gene, and allowing expression to  
 CC reach therapeutically effective levels. The present sequence represents a  
 CC specifically claimed RSV enhancer/promoter nucleotide sequence from the

CC present invention  
 XX Sequence 648 BP; 163 A; 135 C; 179 G; 171 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 100; DB 4; Length 648;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-28;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 60  
 DB 90 CTGCTCCCTGCTTGTGTGTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 149  
 QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
 DB 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189  
 RESULT 3  
 AAV58058  
 ID AAV58058 standard; DNA; 1070 BP.  
 XX AAV58058;  
 AC AAV58058;  
 DT 27-AUG-2003 (revised)  
 DT 11-JAN-1999 (first entry)  
 DE Plasmid CMV-delRT SstI fragment.  
 XX FIV; FIPV; vaccine; reverse transcriptase; diagnosis; therapy; CMV-delRT;  
 KW promoter; cat; ss.  
 XX Human cytomegalovirus.  
 OS feline immunodeficiency virus.  
 OS Chimeric.  
 XX Key Location/Qualifiers  
 FT promoter 8..896  
 FT /tag= a  
 FT /note= "CMV promoter fragment from pcDNA3 (BgIII-KpnII)"  
 FT provirus 918..1070  
 FT /tag= b  
 FT /note= "FIV sequences from primer binding site to SstI  
 FT site"  
 XX WO9840493-A1.  
 XX 17-SEP-1998.  
 XX 10-MAR-1998; 98WO-GB0000715.  
 XX 11-MAR-1997; 97GB-00004977.  
 XX (UNIU ) UNIV GLASGOW.  
 XX Neil JC, Rigby MA, Jarrett JO;  
 XX WPI; 1998-520813/44.  
 XX Protecting, e.g. cats, against feline immunodeficiency virus - by using  
 XX vaccine comprising FIV pol gene containing deletion and/or insertion in  
 XX reverse transcriptase domain.  
 XX Example 3; Fig 4; 66pp; English.  
 XX This is the nucleotide sequence of a SstI fragment of plasmid CMV-delRT,  
 XX in which the immediate-early promoter of human cytomegalovirus replaces  
 XX the 5' long terminal repeat region of feline immunodeficiency virus (FIV)  
 XX clone F14-delRT (see AAV58053). FIV sequences downstream of the SstI site  
 XX are identical to those in F14-delRT. Use of the CMV promoter was designed  
 XX to enhance expression of FIV antigens, and to reduce the risk of  
 XX reversion to a replicating provirus, in tissues after inoculation of DNA.  
 XX Vaccine formulations for FIV-related diseases include a defective feline  
 XX immunodeficiency proviral (FIPV) polynucleotide comprising an in-frame



CC deletion and/or insertion in the reverse transcriptase (RT) region of the  
CC pol gene. Host cells comprising the FIV are capable of producing FIV  
CC proteins, except for functionally competent RT, and thus release non-  
CC infectious FIV viral particles. (Updated on 27-AUG-2003 to correct OS  
CC field.)

XX Sequence 1070 BP; 275 A; 254 C; 268 G; 273 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1070;  
Best Local Similarity 100.0%; Pred. No. 1e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTTGGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTTACA 60  
DB 77 CTGCTCCCTGCTGTGTTGGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTTACA 136

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
DB 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

## RESULT 4

ID ADM41035 standard; DNA; 1506 BP.

XX ADM41035;

XX 17-JUN-2004 (first entry)

XX Fungus nucleotide sequence SEQ ID NO:3.

XX engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX tissue transplantation; human disease study; fungus; gene; ds.

XX Unidentified.

XX WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a fetal non-human mammal host.

XX Disclosure; SEQ ID NO 3; 48pp; English.

XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the fetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for  
XX transplantation, also useful to study human diseases. The present  
XX sequence represents a nucleotide sequence given in the Sequence Listing  
XX of the present invention but not mentioned further within the  
XX specification.

XX Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTTGGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTTACA 60  
DB 81 CTGCTCCCTGCTGTGTTGGAGTCTGCTGAGTGTGCGGAGCAAAATTTAAGCTTACA 140

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100

DB 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

## RESULT 5

ID ADH11349 standard; DNA; 1600 BP.

XX ADH11349;

XX 11-MAR-2004 (first entry)

XX Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
XX cell shape regulator; cell motility regulator; cell migration;  
XX cell behaviour regulator; phenotype; signal transduction pathway;  
XX signal transducing protein; signal integrator protein;  
XX neuronal regeneration; revascularisation; wound healing;  
XX chronic neurodegenerative disease; acute traumatic injury;  
XX fibrotic disease; gene; ds.

XX Unidentified.

XX WO9824810-A2.

XX 11-JUN-1998.

XX 03-DEC-1997; 97WO-EP006956.

XX 04-DEC-1996; 96GB-00025283.

XX (JANC) JANSSEN PHARM NV.

XX Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

XX Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

XX Geysen J, Bogaert TA0B;

XX WPI; 1998-362411/31.

XX P-PSDB; ADH11350.

XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

XX promoting neuronal regeneration, treating chronic neuro-degenerative

XX diseases or acute traumatic injuries.

XX Disclosure; Page 410-411; 479pp; English.

XX The present invention describes a vertebrate protein homologue of an UNC-  
XX 53 protein of Caenorhabditis elegans or a functional equivalent,  
XX derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
XX encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
XX nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
XX the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
XX a transgenic cell, tissue or animal comprising the vector as in (3); (6)  
XX a compound identified as an enhancer or inhibitor of the regulation of  
XX cell shape, motility, or the direction of cell migration for use as a  
XX therapeutic; (7) a method for determination of whether a protein is an  
XX inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
XX motility or the direction of migration by contacting a host cell  
XX expressing a homologue of UNC-53 and determining a change of phenotype;  
XX (8) a method for identification of vertebrate homologues of C. elegans  
XX unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
XX a DNA library; and (9) a method for identification of a protein which is

CC active in the signal transduction pathway of a cell of which a vertebrate  
CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
CC antibody/homologue complex; and (iii) analysing such a complex to  
CC identify any non-antibody protein bound to the complex. UNC-53 is a  
CC signal transducing or signal integrator protein involved in controlling  
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate  
CC homologues of UNC-53 can be used to promote neuronal regeneration.  
CC revascularisation or wound healing, to treat chronic neurodegenerative  
CC diseases or acute traumatic injuries or fibrotic diseases. The present  
CC sequence is used in the exemplification of the present invention.  
XX  
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGCTCCCTGCTGTGTGCTGAGGTGCTGAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 60  
DB 81 CTGCTCCCTGCTGTGTGCTGAGGTGCTGAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 140  
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 100  
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 180

RESULT 6  
ADM41037  
ID ADM41037 standard; DNA; 1782 BP.  
XX  
AC ADM41037;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.  
XX  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
XX  
OS Cytomegalovirus.  
XX  
FN WO2004027029-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 17-SEP-2003; 2003WO-US029251.  
XX  
PR 19-SEP-2002; 2002US-0411790P.  
XX  
PA (XIME-) XIMEREX INC.  
XX  
PI Beschornor WE, Sosa CE, Thompson SC;  
XX  
WPI; 2004-295402/27.  
XX  
PT Engrafting foreign replacement cells within a fetal non-human mammal,  
PT useful in producing chimeric mammals, comprises selectively destroying  
PT native cells in a tissue of a fetal non-human mammal host.  
XX  
PS Disclosure; SEQ ID NO 5; 48pp; English.  
XX  
CC The present invention describes a method for engrafting foreign  
CC replacement cells within a fetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a fetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the fetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.  
XX  
SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 12; Length 1782;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGCTCCCTGCTGTGTGCTGAGGTGCTGAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 60  
DB 81 CTGCTCCCTGCTGTGTGCTGAGGTGCTGAGTAGTGCAGCGAGCAAAATTTAAAGCTACA 140  
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 100  
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 180  
RESULT 7  
ADM41034  
ID ADM41034 standard; DNA; 2241 BP.  
XX  
AC ADM41034;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Human nucleotide sequence SEQ ID NO:2.  
XX  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO2004027029-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 17-SEP-2003; 2003WO-US029251.  
XX  
PR 19-SEP-2002; 2002US-0411790P.  
XX  
PA (XIME-) XIMEREX INC.  
XX  
PI Beschornor WE, Sosa CE, Thompson SC;  
XX  
WPI; 2004-295402/27.  
XX  
PT Engrafting foreign replacement cells within a fetal non-human mammal,  
PT useful in producing chimeric mammals, comprises selectively destroying  
PT native cells in a tissue of a fetal non-human mammal host.  
XX  
PS Disclosure; SEQ ID NO 2; 48pp; English.  
XX  
CC The present invention describes a method for engrafting foreign  
CC replacement cells within a fetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a fetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the fetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.  
XX  
SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;  
Best Local Similarity 100.0%; Pred. No. 1.3e-27; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0;

QY 1 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60  
DB 81 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 140

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 8  
ID ABZ23249 standard; DNA; 2245 BP.  
AC ABZ23249;  
DT 24-MAR-2003 (first entry)  
DE Lac repressor operated p21-expression cassette and RSV-LTR promoter.  
KW Lac repressor; p21; RSV; LTR promoter; cell cycle inhibitor protein;  
KW protein production; anchorage-independent producer cell line; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT promoter 1..563  
FT /tag= a  
FT /note= "RSV-LTR promoter"  
FT intron 564..1051  
FT /tag= b  
FT /note= "SV40 small t antigen intron"  
FT misc\_feature 1052..1907  
FT /tag= c  
FT /note= "p21 coding sequence"  
FT polyA\_signal 1908..2245  
FT /tag= d  
FT /note= "thymidine kinase polyA site"

WO200299100-A2.  
12-DEC-2002.  
03-JUN-2002; 2002WO-EP006054.  
PR 01-JUN-2001; 2001GB-00013318.  
XX (LONZ ) LONZA BIOLOGICS PLC.  
XX  
XX Al-Rubeai M, Shuttleworth J;  
XX WPI; 2003-148669/14.  
XX  
XX Producing recombinant protein, particularly for maximizing or enhancing  
XX e.g. therapeutic protein production, by co-expressing protein with  
XX recombinant cell cycle inhibitor protein (p21) in producer cell line.  
XX  
XX Example 1; Page 15-16; 33pp; English.  
XX  
XX The present sequence represents a lac repressor operated p21-expression  
XX cassette comprising the Rous sarcoma virus (RSV)-LTR promoter. p21 is a  
XX cell cycle inhibitor protein. The present sequence is used to produce  
XX vectors for use in the method of the invention. The specification  
XX describes a method for producing a protein, preferably a recombinant  
XX protein, in a mammalian anchorage-independent producer cell line. The  
XX method comprises co-expressing with the protein in the producer cell line  
XX a recombinant cell cycle inhibitor protein (preferably p21). The method  
XX is useful for producing a recombinant protein in a producer cell line.  
XX This is particularly useful for maximizing or enhancing the production of  
XX e.g. therapeutic proteins at an industrial scale

XX SQ Sequence 2245 BP; 532 A; 555 C; 625 G; 533 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 8; Length 2245;  
Best Local Similarity 100.0%; Pred. No. 1.3e-27; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0;

QY 1 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60  
DB 46 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 105

QY 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100  
DB 106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145

RESULT 9  
ID ADM41036 standard; DNA; 2294 BP.  
AC ADM41036;  
DT 17-JUN-2004 (first entry)  
DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.  
KW engrafting foreign replacement cell; implanting foreign replacement cell;  
KW growth; differentiation; drug development; vaccine development;  
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
OS Cytomegalovirus.  
FN WO2004027029-A2.  
PD 01-APR-2004.  
PF 17-SEP-2003; 2003WO-US029251.  
PR 19-SEP-2002; 2002US-0411790P.  
XX (XIME-) XIMEREX INC.  
XX Beschornier WE, Sosa CE, Thompson SC;  
XX WPI; 2004-295402/27.  
XX  
XX Engrafting foreign replacement cells within a fetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a fetal non-human mammal host.  
XX  
XX Disclosure; SEQ ID NO 4; 48pp; English.  
XX  
XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the fetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for  
XX transplantation, also useful to study human diseases. The present  
XX sequence represents a nucleotide sequence given in the Sequence Listing  
XX of the present invention but not mentioned further within the  
XX specification.

XX SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 12; Length 2294;  
Best Local Similarity 100.0%; Pred. No. 1.3e-27; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0;

```
Oy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
    |||||||
Db 81 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 140

Oy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
    |||||||
Db 141 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 180

RESULT 10
AAD02037
ID AAD02037 standard; DNA; 2426 BP.
XX
AC AAD02037;
XX
DT 11-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
DE Plasmid pNG3/RC/CPG2-Thy1 comprising CPG2 DNA with rat thy1 gene.
XX
KW Carboxypeptidase G2; CPG2; gene directed enzyme prodrg therapy; GDEPT;
KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
KW plasmid; ds.
XX
OS Rattus sp.
OS Bacteria.
OS Chimeric.
XX
PN WO200066752-A2.
XX
PD 09-NOV-2000.
XX
PF 28-APR-2000; 2000WO-GB001640.
XX
PR 01-MAY-1999; 99GB-00010077.
XX
PA (ASTR ) ASTRAZENECA AB.
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Castro MG, Emery SC, Lowenstein PR;
XX
WPI; 2001-015983/02.
XX
Gene directed enzyme prodrg therapy using post translational
glycosylphosphatidylinositol addition to prodrg activating enzyme to
enable anchorage of enzyme at cell surface for cancer therapy.
XX
Example 1e; Page 59-60; 60pp; English.
XX
The present invention relates to a gene directed enzyme prodrg therapy
(GDEPT) using post translational glycosylphosphatidylinositol (GPI)
addition to a prodrg activating enzyme which enables anchorage of the
enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
prodrug activating enzyme. The invention also relates to an expression
vector for expression of a GPI enzyme hybrid capable of anchorage to the
surface of a mammalian cell. The expression vector comprise
polynucleotide sequences encoding a signal peptide, an enzyme capable of
activating a prodrg, and a post-translational GPI addition motif. The
expression vector is useful in the manufacture of a medicament for cancer
therapy in a mammalian host. The present DNA sequence is a plasmid
pNG3/RC/CPG2-Thy1 comprising CPG2 nucleic acid sequence with the last
standardise OS field)
SQ Sequence 2426 BP; 557 A; 705 C; 668 G; 495 T; 0 U; 1 Other;
```

```
Query Match 100.0%; Score 100; DB 4; Length 2426;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
    |||||||
Db 69 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 128
```

```
Oy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
    |||||||
Db 129 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 168

RESULT 11
AAD02036
ID AAD02036 standard; DNA; 2427 BP.
XX
AC AAD02036;
XX
DT 11-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
DE Plasmid pNG3/RC/CPG2 (Q3) -Thy1 comprising CPG2 variant with rat thy1 gene.
XX
KW Carboxypeptidase G2; CPG2; gene directed enzyme prodrg therapy; GDEPT;
KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
KW CPG2 (Q3) variant; plasmid; ds.
XX
OS Rattus sp.
OS Bacteria.
OS Chimeric.
XX
PN WO200066752-A2.
XX
PD 09-NOV-2000.
XX
PF 28-APR-2000; 2000WO-GB001640.
XX
PR 01-MAY-1999; 99GB-00010077.
XX
PA (ASTR ) ASTRAZENECA AB.
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Castro MG, Emery SC, Lowenstein PR;
XX
WPI; 2001-015983/02.
XX
Gene directed enzyme prodrg therapy using post translational
glycosylphosphatidylinositol addition to prodrg activating enzyme to
enable anchorage of enzyme at cell surface for cancer therapy.
XX
Example 1e; Page 59; 60pp; English.
XX
The present invention relates to a gene directed enzyme prodrg therapy
(GDEPT) using post translational glycosylphosphatidylinositol (GPI)
addition to a prodrg activating enzyme which enables anchorage of the
enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
prodrug activating enzyme. The invention also relates to an expression
vector for expression of a GPI enzyme hybrid capable of anchorage to the
surface of a mammalian cell. The expression vector comprise
polynucleotide sequences encoding a signal peptide, an enzyme capable of
activating a prodrg, and a post-translational GPI addition motif. The
expression vector is useful in the manufacture of a medicament for cancer
therapy in a mammalian host. The present DNA sequence is a plasmid
pNG3/RC/CPG2 (Q3) comprising CPG2 variant (Q3) and the last exon of
rat Thy-1 at the 3' end. (Updated on 11-SEP-2003 to standardise OS field)
SQ Sequence 2427 BP; 555 A; 706 C; 670 G; 495 T; 0 U; 1 Other;
```

```
Query Match 100.0%; Score 100; DB 4; Length 2427;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
    |||||||
Db 70 CTGCTCCCTGCTGTGTGTTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 129

Oy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
    |||||||
Db 130 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 169
```

RESULT 12  
 AAT62937  
 ID AAT62937 standard; DNA; 3400 BP.  
 XX AC AAT62937;  
 XX DT 17-OCT-2003 (revised)  
 XX DT 16-JUN-1997 (first entry)  
 XX DE 3F4 human G2/G4 chimeric antibody expression plasmid insert.  
 XX KW Xenotransplantation; graft rejection; cell interaction; pig;  
 XX KW vascular cell adhesion molecule; VCAM; monoclonal antibody;  
 XX KW chimeric antibody; diagnosis; ss.  
 XX OS Homo; sapiens.  
 XX OS Mus sp.  
 XX OS Chimeric.  
 XX FH Key Location/Qualifiers  
 FT exon 903..1055  
 FT FT /\*tag= a  
 FT intron 1056..1285  
 FT FT /\*tag= b  
 FT exon 1286..2055  
 FT FT /\*tag= c  
 FT FT /codon\_start= 1350  
 FT intron 2056..2447  
 FT FT /\*tag= d  
 FT exon 2448..2483  
 FT FT /\*tag= e  
 FT intron 2484..2601  
 FT FT /\*tag= f  
 FT exon 2602..2928  
 FT FT /\*tag= g  
 FT intron 2929..3025  
 FT FT /\*tag= h  
 FT exon 3026..3348  
 FT FT /\*tag= i  
 XX WO9711971-Al.  
 PN 03-APR-1997.  
 PD 27-SEP-1996; 96WO-US015575.  
 XX PR 28-SEP-1995; 95US-0004489P.  
 XX PR 26-SEP-1996; 96US-00004489.  
 XX PA (ALEX-) ALEXION PHARM INC.  
 XX PI Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;  
 XX DR WPI; 1997-212855/19.  
 XX DR P-PSDB; AAW14940.  
 XX PT Antibodies binding to porcine but not human cell interaction proteins -  
 PT useful to treat and assay for rejection of xenografted porcine organs,  
 PT tissues or cells.  
 XX PS Disclosure; Page 58-61; 105pp; English.  
 XX CC A DNA sequence (AAT62937) comprises a 3F4 human G2/G4 (see also AAT62936)  
 CC chimeric antibody expression plasmid insert sequence. The chimeric  
 CC antibody (AAW14940) is specific for porcine vascular cell adhesion  
 CC molecule (VCAM) and is useful for diagnosing human rejection of porcine  
 CC xenotransplants and for improving xenotransplantation of porcine cells,  
 CC tissues and organs into human recipients. (Updated on 17-OCT-2003 to  
 CC standardise OS field)  
 XX SQ Sequence 3400 BP; 759 A; 1012 C; 909 G; 720 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 3400;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-27;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGCTCCCTGCTTGTGTGGAGTGGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
 |||||||  
 Db 148 CTGCTCCCTGCTTGTGTGGAGTGGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 207  
 |||||||  
 QY 61 ACAAGGCAAGCTTGAACCGACAATTGCATGAAGAATCTGC 100  
 |||||||  
 Db 208 ACAAGGCAAGCTTGAACCGACAATTGCATGAAGAATCTGC 247  
 |||||||

RESULT 13  
 AAT62932  
 ID AAT62932 standard; DNA; 3400 BP.  
 XX AC AAT62932;  
 XX DT 17-OCT-2003 (revised)  
 XX DT 16-JUN-1997 (first entry)  
 XX DE 2A2 human G2/G4 chimeric antibody expression plasmid insert.  
 XX KW Xenotransplantation; graft rejection; cell interaction; pig;  
 XX KW vascular cell adhesion molecule; VCAM; monoclonal antibody;  
 XX KW chimeric antibody; diagnosis; ss.  
 XX OS Homo; sapiens.  
 XX OS Mus sp.  
 XX OS Chimeric.  
 XX FH Key Location/Qualifiers  
 FT exon 903..1055  
 FT FT /\*tag= a  
 FT intron 1056..1285  
 FT FT /\*tag= b  
 FT exon 1286..2020  
 FT FT /\*tag= c  
 FT FT /codon\_start= 1318  
 FT intron 2021..2412  
 FT FT /\*tag= d  
 FT exon 2413..2448  
 FT FT /\*tag= e  
 FT intron 2449..2566  
 FT FT /\*tag= f  
 FT exon 2567..2983  
 FT FT /\*tag= g  
 FT intron 2984..2990  
 FT FT /\*tag= h  
 FT exon 2991..3313  
 FT FT /\*tag= i  
 XX WO9711971-Al.  
 PN 03-APR-1997.  
 PD 27-SEP-1996; 96WO-US015575.  
 XX PR 28-SEP-1995; 95US-0004489P.  
 XX PR 26-SEP-1996; 96US-00004489.  
 XX PA (ALEX-) ALEXION PHARM INC.  
 XX PI Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;  
 XX DR WPI; 1997-212855/19.  
 XX DR P-PSDB; AAW14934.  
 XX PT Antibodies binding to porcine but not human cell interaction proteins -  
 PT useful to treat and assay for rejection of xenografted porcine organs,  
 PT tissues or cells.



PT Selectable retroviral packaging cell lines and expression constructs -  
PT comprise selectable gene downstream of gene of interest, are selectable  
PT due to the in-efficiency associated with translation re-initiation.

PS Claim 23; Fig 13; 79pp; English.

XX  
XX  
CC This sequence represents the recombinant expression plasmid CMV10A. This  
CC sequence is a packaging-deficient construct having a viral env gene (in  
CC this case from moloney murine leukaemia virus under hCMV promoter  
CC control) and a selectable marker (SM). It is an example of a recombinant  
CC expression vector (REV) of the invention, used to create a packaging cell  
CC line. The REVs of the invention comprise a gene of interest (GOI) and a  
CC SM gene. The SM gene is arranged downstream of the GOI and a GOI  
CC associated stop codon is spaced from a start codon of the SM gene to  
CC ensure that the SM protein is expressed as a result of translation  
CC reinitiation. The cell lines are transformed with two REV's, both are  
CC replication deficient, one contains the viral gag-pol gene, the other the  
CC viral env gene. By using helper constructs, such as the REV's, which are  
CC directly selectable and which provide for high expression of the viral  
CC gene, high titre retroviral vectors may be obtained. The packaging cell  
CC lines are useful for gene therapy. Prior packaging cell lines using full  
CC length retroviral genomes as helper genomes were isolated by  
CC cotransfecting them with plasmids encoding selectable markers. However,  
CC the helper functions can be lost during the passages of the cells in  
CC culture and the current packaging systems provide limited titres of  
CC infectious retroviral vectors. Co-transfection with a plasmid encoding a  
CC SM does not directly select the best gag-pol-env-expressing cells. The  
CC new retroviral packaging cell lines overcome these problems

XX SQ Sequence 3925 BP; 963 A; 1001 C; 959 G; 998 T; 0 U; 4 Other;

Query Match 100.0%; Score 100; DB 2; Length 3925;  
Best Local Similarity 100.0%; Pred. No. 1.6e-27;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGCTCCCTGCTTGTGTGAGTCGCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60  
Db 70 CTGCTCCCTGCTTGTGTGAGTCGCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 129  
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAACTCTGC 100  
Db 130 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAACTCTGC 169

Search completed: July 14, 2005, 07:01:48  
Job time : 141.038 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 ctgtccctctgtgtgtt.....caattgcatgaagaattctgc 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	602	8	B67169 CP00047A Cp
2	30.4	30.4	829	4	BI333630 602997459
3	30.2	30.2	823	6	CD655614 AGENCOURT
4	29.8	29.8	1165	8	CC242469 CH261-11P
5	29.4	29.4	401	6	CB387202 OSTF076E6
6	29.4	29.4	531	5	BQ310441 MR0-BT450
7	29.4	29.4	754	8	AQ946479 Sheared D
8	29.2	29.2	426	9	CC888514 SALK 1519
9	29	29.0	657	7	CK086063 RG11-C07
10	29	29.0	877	9	AL225351 Tetradon
11	28.6	28.6	340	7	F32722 HSPD25699 H
12	28.6	28.6	408	1	AA962465 oo91e05.8
13	28.6	28.6	436	7	CN386744 328755673
14	28.6	28.6	514	6	CB161201 K-EST0221
15	28.6	28.6	530	6	CB161182 K-EST0220
16	28.6	28.6	534	2	AW500392 UI-HF-BN0
17	28.6	28.6	550	1	AA984313 am83h04.8
18	28.6	28.6	555	7	CR537056 DKFZp459D
19	28.6	28.6	583	7	CN386728 170005326
20	28.6	28.6	585	4	BG993413 MR3-HT099
21	28.6	28.6	625	4	BI113747 602860946
22	28.6	28.6	626	7	CN386724 170006000
23	28.6	28.6	641	2	AW955076 EST367146
24	28.6	28.6	651	7	CN386730 170005316

C 25	28.6	28.6	680	7	CN386731 170005313
C 26	28.6	28.6	682	7	CN386746 170005999
C 27	28.6	28.6	687	7	CN386717 170004551
C 28	28.6	28.6	710	2	BF309673 601891808
C 29	28.6	28.6	724	7	CN386780 170005318
C 30	28.6	28.6	770	6	CD654097 AGENCOURT
C 31	28.6	28.6	777	7	CN386742 170005314
C 32	28.6	28.6	801	6	CD656906 AGENCOURT
C 33	28.6	28.6	814	4	BG820298 602782110
C 34	28.6	28.6	823	6	CB996419 AGENCOURT
C 35	28.6	28.6	827	4	BG387788 602412672
C 36	28.6	28.6	830	6	CD643822 AGENCOURT
C 37	28.6	28.6	842	5	BU170446 AGENCOURT
C 38	28.6	28.6	851	6	CB993607 AGENCOURT
C 39	28.6	28.6	852	5	BQ222104 AGENCOURT
C 40	28.6	28.6	856	6	CD656102 AGENCOURT
C 41	28.6	28.6	909	5	BU189860 AGENCOURT
C 42	28.6	28.6	1025	4	BM477450 AGENCOURT
C 43	28.4	28.4	579	6	CB239722 RSH15G08
C 44	28.4	28.4	733	7	CF667137 RSH15G08
C 45	28.4	28.4	1015	6	BY703355 BY703355

#### ALIGNMENTS

RESULT 1  
B67169  
LOCUS CP00047A CpioWagDNA2 Cryptosporidium parvum genomic, genomic survey  
DEFINITION 602 bp DNA linear GSS 12-MAY-2000  
sequence.  
ACCESSION B67169  
VERSION B67169.1 GI:2642750  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum  
ORGANISM Cryptosporidium parvum  
REFERENCE 1 (bases 1 to 602)  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an  
expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
MEDLINE 20183851  
PUBMED 10717299  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San  
Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malariaditaa.ucsf.edu  
Submitted sequence has been edited to remove vector sequences 5' to  
the insert, to correct miscalled bases and assign uncalled (N)  
bases throughout the sequence, and to terminate when base-calling  
became ambiguous.  
Seq primer: T7  
Class: shotgun  
High quality sequence stop: 602.  
Location/Qualifiers  
1. .602  
/organism="Cryptosporidium parvum"  
/mol\_type="genomic DNA"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/lab\_host="E. coli XL2 Blue MRF"  
/clone\_lib="CpioWagDNA2"  
/note="Vector: PCR-Script Amp SK+; Site 1: SrfI; C. parvum  
(IOWA isolate) genomic DNA was hydrodynamically sheared  
to produce fragments having a tight size distribution  
between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford  
DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

#### ORIGIN

Query Match 100.0%; Score 100; DB 8; Length 602;  
 Best Local Similarity 100.0%; Pred. No. 8e-24;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCTGCTGTTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60  
 |||||  
 Db 41 CTGCTCCTGCTGTTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 100  
 |||||

Qy 61 ACAAGCAAGCTTGACGACCAATGTCATGAAGATCTGC 100  
 |||||  
 Db 101 ACAAGCAAGCTTGACGACCAATGTCATGAAGATCTGC 140  
 |||||

#### RESULT 2

B1333630/c  
 LOCUS 602997459F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5139470 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1333630  
 VERSION B1333630.1 GI:15018287  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 829)  
 NIH-MGC http://mhc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Sequencing by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM11343 row: f column: 15  
 High quality sequence stop: 788.  
 Location/Qualifiers  
 1..829  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5139470"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."

#### FEATURES

source  
 1..829  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5139470"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."

#### ORIGIN

Query Match 30.4%; Score 30.4; DB 4; Length 829;  
 Best Local Similarity 67.2%; Pred. No. 15;  
 Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 19 TTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACAACAGCGAGCTTGACC 78  
 |||||

Db 217 TTGGGCGTCCAGAAATTGTTGGTGAGCAAACTTCAAGTTGCTGCTGGGAAGTCTTGACT 158  
 Qy 79 GACA 82  
 |||||  
 Db 157 GACA 154

RESULT 3  
 CD655614/c  
 LOCUS 823 bp mRNA linear EST 18-JUN-2003  
 DEFINITION AGENCOURT 14551032 NIA Human H1 Embryonic Stem Cell cDNA Library (long) Homo sapiens cDNA clone IMAGE:30424285 5', mRNA sequence.  
 ACCESSION CD655614  
 VERSION CD655614.1 GI:31896113  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 823)  
 NIH-MGC http://mhc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
 cDNA Library Preparation: Yulan Piao and Minoru Ko  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM506 row: k column: 14  
 High quality sequence stop: 640.  
 Location/Qualifiers  
 1..823  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30424285"  
 /tissue\_type="Embryonic Stem cells"  
 /cell\_line="WA01"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"  
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MSF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCTTTT-TTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker IL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase

003

LOCUS  
DEFINITION  
ACCESSION  
VERSION

DEFINITION MKO-BT4502-Z20601-202-B02 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BQ310441  
 VERSION BQ310441.1 GI:20853032



Salk Institute Genomic Analysis Laboratory (SIGnAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of  
TDNA. This sequence lies within an annotated intron of At1g03960.  
Class: TDNA tagged.

#### FEATURES

Location/Qualifiers  
1..426  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="SALK151964.54.50.x"  
/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
/note="PCR was performed on Arabidopsis thaliana lines  
each of which contains one or more TDNA insertion  
elements. The resultant fragment for each line was  
directly sequenced to determine the genomic sequence at  
the site of insertion. Details of the protocols used can  
be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

#### ORIGIN

Query Match 29.2%; Score 29.2; DB 9; Length 426;  
Best Local Similarity 64.2%; Pred. No. 35;  
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
Qy 30 TGAGTAGTCGCGAGCAAAATTTAGCTACACAGGAGGCTTGACCGACAATTCGAT 89  
Db 76 TTATTAGTTGGTGTTCAGANTTTAGCATCATCAAGACTTGACCTACAGAACAT 17  
Qy 90 GAACAAT 96  
Db 16 GAAAAAT 10

RESULT 9  
CK086063/c  
LOCUS CK086063 657 bp mRNA linear EST 01-DEC-2003  
DEFINITION RG11\_C07 Cucumber leaf Cucumis sativus cDNA, mRNA sequence.  
ACCESSION CK086063  
VERSION CK086063.1 GI:38571123  
KEYWORDS EST.  
SOURCE Cucumis sativus (cucumber)  
ORGANISM Cucumis sativus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
1 (bases 1 to 657)  
Grumet, R. and McGrath, M.  
Development of genomic tools for cucumber (Cucumis sativus L.)  
Unpublished (2003)  
Contact: Rebecca Grumet  
Rebecca Grumet  
Michigan State University  
Horticulture Department, Michigan State University, East Lansing,  
MI 48824, USA  
Tel: 517 353 0890  
Fax: 517 355 5191 x431  
Email: grumet@msu.edu  
Plate: RG11 row: C column: 07.  
Location/Qualifiers  
1..657  
/organism="Cucumis sativus"  
/mol\_type="mRNA"  
/strain="Straight 8"  
/db\_xref="taxon:3659"  
/sex="monoecious"  
/clone\_lib="Cucumber leaf"  
/note="Vector: pAD-GAL4; Site\_1: EcoRI; Site\_2: XhoI"

#### FEATURES

Location/Qualifiers  
1..657  
/organism="Cucumis sativus"  
/mol\_type="mRNA"  
/strain="Straight 8"  
/db\_xref="taxon:3659"  
/sex="monoecious"  
/clone\_lib="Cucumber leaf"  
/note="Vector: pAD-GAL4; Site\_1: EcoRI; Site\_2: XhoI"

#### ORIGIN

Query Match 29.0%; Score 29; DB 7; Length 657;  
Best Local Similarity 63.8%; Pred. No. 44;  
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 31 GAGTAGTCGCGAGCAAAATTTAGCTACACAGGAGGCTTGACCGACAATTCGATG 90  
Db 267 GAGGAGTGCAGGAAGCAAACTGAAGCCAGAATGAAGAGAGGCTGGACTTCGAAGGTCAAG 208  
Qy 91 AAGAATCTG 99  
Db 207 AAAAATTTG 199

#### RESULT 10

CNS032VI 877 bp DNA linear GSS 01-SEP-2000  
LOCUS CNS032VI/c  
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone  
sequence.  
ACCESSION AL225351  
VERSION AL225351.1 GI:7884242  
KEYWORDS GSS; genome survey sequence.  
SOURCE Tetraodon nigroviridis  
ORGANISM Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontoidea; Tetraodontidae; Tetraodon.

#### REFERENCE

1 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Barnot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
PUBMED 10835645

#### REFERENCE

1 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Barnot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
PUBMED 10835645

#### REFERENCE

1 Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)  
20359837  
PUBMED 10899143

#### REFERENCE

3 (bases 1 to 877)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 131 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

#### COMMENT

This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetraodon>.

#### FEATURES

Location/Qualifiers  
1..877  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="207F02"  
/clone\_lib="G"  
/notes="Genoscope sequence ID : COAG207DC01SP1-end :  
PUC-Ori"

#### ORIGIN

Query Match 29.0%; Score 29; DB 9; Length 877;  
Best Local Similarity 63.8%; Pred. No. 47;  
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 14 GTGTGTGGAGTCGCTGAGTAGTCGCGAGCAAAATTTAAGCTACACAGGCAAGGCT 73

```

Db 199 GGGTCCTGGGGGCGCTGTCGGTGCTGAAGCATAAACATCTACAACTGGAAGTTA 140
Qy 74 TGACCGACA 82
Db 139 TGAAGAGA 131

RESULT 11
F32722/c
LOCUS F32722 340 bp mRNA linear EST 13-MAY-1999
DEFINITION HSPD25699 HM3 Homo sapiens cDNA clone s3000037G06, mRNA sequence.
ACCESSION F32722
VERSION F32722.1 GI:4818348
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 340)
AUTHORS Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,
Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
PubMed 8681137
COMMENT Contact: Valle G.
CRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://group.bio.unipd.it.
Location/Qualifiers
FEATURES
source
1..340
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="s3000037G06"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/clone_lib="HM3"
/notes="vector: pcDNAII (Invitrogen); Site 1: BstXI;
Site 2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGCGCTTTT-TTTT-TTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pcDNAII vector."
ORIGIN
Query Match 28.6%; Score 28.6; DB 7; Length 340;
Best Local Similarity 61.3%; Pred. No. 55;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCCTGAGTGTGCGGAGCAAAATTTAAGCTACAAAGGCGAGGCTTGACCGAC 81
Db 125 GAGCTAACTGAATTGTATGGAGCAGCATTTAACATATTCCTAGTCAAGGACGAGTGGG 66
Qy 82 AATTGCATGAAGAT 96
Db 65 AAGTAAGTGAAGAT 51

RESULT 12
AA962465
LOCUS AA962465 408 bp mRNA linear EST 07-JUL-1998
DEFINITION NC091e05.81 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1573568 3',

```

```

mRNA sequence.
AA962465
AA962465.1 GI:3134629
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1045 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers
FEATURES
source
1..408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1573568"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid5"
/notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AATGGAAGAAATCGCGCGCAATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match 28.6%; Score 28.6; DB 1; Length 408;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGTGAGTGTGCGGAGCAAAATTTAAGCTACAAAGGCGAGGCTTGACCGAC 81
Db 243 GAGCTAACTGAATTGTATGGAGCAGCATTTAACATATTCCTAGTCAAGGACGAGTGGG 302
Qy 82 AATTGCATGAAGAT 96
Db 303 AAGTAAGTGAAGAT 317

RESULT 13
CN386744/c
LOCUS CN386744 436 bp mRNA linear EST 16-MAY-2004
DEFINITION 328755673 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN386744
VERSION CN386744.1 GI:47374339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 436)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J and Stanton, L.W.

```

```

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT    Contact: Brandenberger R
            Regenerative Medicine
            Genon Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@genon.com
FEATURES   Insert Length: 436 Std Error: 0.00.
            Location/Qualifiers
            source
            1..436
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, embryoid bodies
            derived from H1, H7 and H9 cells"
            /clone_lib="GRN_EB"
            /note="oligo dt primed, full-length enriched cDNA library
            from embryoid body outgrowths derived from hES cell lines
            H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
            conditions."

ORIGIN
Query Match      28.6%; Score 28.6; DB 7; Length 436;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 133 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACGAGTGGG 74

Qy 82 AATTGCATGAAGAAT 96
Db 73 AAGTAAGTGAAGAAT 59

RESULT 14
LOCUS      CB161201
DEFINITION K-EST0221011 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-H03
5', mRNA sequence.
ACCESSION  CB161201
VERSION     CB161201.1 GI:28147327
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 514)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16 row: H column: 03
            High quality sequence stop: 514.
            Location/Qualifiers
            source
            1..514
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-H03"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT    Contact: Brandenberger R
            Regenerative Medicine
            Genon Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@genon.com
FEATURES   Insert Length: 436 Std Error: 0.00.
            Location/Qualifiers
            source
            1..436
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-H03"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 514;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACGAGTGGG 320

Qy 82 AATTGCATGAAGAAT 96
Db 321 AAGTAAGTGAAGAAT 335

RESULT 15
LOCUS      CB161182
DEFINITION K-EST0220988 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-F04
5', mRNA sequence.
ACCESSION  CB161182
VERSION     CB161182.1 GI:28147308
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 530)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16 row: F column: 04
            High quality sequence stop: 530.
            Location/Qualifiers
            source
            1..530
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-F04"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"
            /clone_lib="L18POOL1n1"
            /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
            Site 2: NotI; The library was contributed by the Soares
            laboratory and it was constructed as described by Bonaldo,
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9): 791-806. RNA was prepared from harvested cell
            culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 530;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACGAGTGGG 320

```

```

/clone_lib="L18POOL1n1"
/note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 514;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACGAGTGGG 320

Qy 82 AATTGCATGAAGAAT 96
Db 321 AAGTAAGTGAAGAAT 335

RESULT 15
LOCUS      CB161182
DEFINITION K-EST0220988 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-F04
5', mRNA sequence.
ACCESSION  CB161182
VERSION     CB161182.1 GI:28147308
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 530)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16 row: F column: 04
            High quality sequence stop: 530.
            Location/Qualifiers
            source
            1..530
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-F04"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"
            /clone_lib="L18POOL1n1"
            /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
            Site 2: NotI; The library was contributed by the Soares
            laboratory and it was constructed as described by Bonaldo,
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9): 791-806. RNA was prepared from harvested cell
            culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 530;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACGAGTGGG 320

```

Qy 82 AATTGCATGAAGAT 96  
|||  
Db 321 AAGTAAGTGAAGAT 335  
|||

Search completed: July 14, 2005, 23:23:20  
Job time : 952.146 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_4192\_4292  
Perfect score: 101  
Sequence: 1 tgccttgccgcgtgctggac.....ccatgggtgctgcctgtgt 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	101	100.0	609	6	AX937037	AX937037 Sequence
2	101	100.0	609	6	AX952091	AX952091 Sequence
3	101	100.0	1973	14	OHVSAGA	M15954 Woodchuck h
4	101	100.0	2649	14	AF410858	AF410858 Woodchuck h
5	101	100.0	2655	14	AF410857	AF410857 Woodchuck h
6	101	100.0	2655	14	AF410859	AF410859 Woodchuck h
7	101	100.0	2655	14	AF410860	AF410860 Woodchuck h
8	101	100.0	2655	14	AF410861	AF410861 Woodchuck h
9	101	100.0	3320	14	OHVGB	M11082 Woodchuck h
10	101	100.0	3323	14	OHVGC	M19183 Woodchuck h
11	101	100.0	3323	14	OHVCGD	M18752 Woodchuck h
12	101	100.0	3323	14	OHVHEPBA	J04514 Woodchuck h
13	101	100.0	7207	12	AY468486	AY468486 Lentivira
14	101	100.0	7515	6	AX663053	AX663053 Sequence
15	101	100.0	7808	10	MAR87	M60766 Woodchuck h
16	101	100.0	8484	6	BD268253	BD268253 Adenoviru
17	101	100.0	10469	6	AX641836	AX641836 Sequence
18	99.4	98.4	2655	14	AF410855	AF410855 Woodchuck h
19	99.4	98.4	2655	14	AF410856	AF410856 Woodchuck h

20	96.2	95.2	3308	14	OHVSURCOR	M90520 Woodchuck h
21	94.6	93.7	592	6	AR136166	AR136166 Sequence
22	94.6	93.7	592	6	AR168222	AR168222 Sequence
23	94.6	93.7	592	6	AR177246	AR177246 Sequence
24	94.6	93.7	3308	6	CO818826	CO818826 Sequence
25	94.6	93.7	3308	14	AY334075	AY334075 Woodchuck
26	94.6	93.7	3308	14	AY334076	AY334076 Woodchuck
27	94.6	93.7	3308	14	AY334077	AY334077 Woodchuck
28	94.6	93.7	3308	14	AY628095	AY628095 Woodchuck
29	94.6	93.7	3308	14	AY628096	AY628096 Woodchuck
30	94.6	93.7	3308	14	AY628097	AY628097 Woodchuck
31	94.6	93.7	3308	14	AY628098	AY628098 Woodchuck
32	94.6	93.7	3308	14	AY628099	AY628099 Woodchuck
33	94.6	93.7	3308	14	AY628100	AY628100 Woodchuck
34	94.6	93.7	3308	14	OHVCGA	J02442 Woodchuck h
35	94.6	93.7	6893	6	AX823860	AX823860 Sequence
36	91.4	90.5	5617	6	AX384541	AX384541 Sequence
37	91.4	90.5	5691	6	AX359937	AX359937 Sequence
38	91.4	90.5	5691	6	AX382151	AX382151 Sequence
39	91.4	90.5	5711	6	AX359934	AX359934 Sequence
40	91.4	90.5	5711	6	AX382148	AX382148 Sequence
41	91.4	90.5	5731	6	AX384542	AX384542 Sequence
42	91.4	90.5	5732	6	AX359932	AX359932 Sequence
43	91.4	90.5	5732	6	AX382146	AX382146 Sequence
44	91.4	90.5	6026	6	AX384539	AX384539 Sequence
45	91.4	90.5	6140	6	AX384540	AX384540 Sequence

#### ALIGNMENTS

RESULT 1

AX937037

LOCUS

AX937037

DEFINITION

Sequence 1 from Patent EP1361277.

ACCESSION

AX937037

VERSION

AX937037.1

KEYWORDS

GI:40713229

SOURCE

synthetic construct

ORGANISM

synthetic construct

REFERENCE

1

AUTHORS

Mallet, J., Brun, S., Dufour, N. and Faucon-Biguet, N.

TITLE

Optimization of transgene expression in mammalian cells

JOURNAL

Patent: EP 1361277-A 1 12-NOV-2003;

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR); Biovectys

(FR)

FEATURES

source

Location/Qualifiers

1..609

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Description of Artificial Sequence: WPRE region."

ORIGIN

Query Match

Best Local Similarity

Matches

100.0%; Score 101.; DB 6; Length 609;

100.0%; Pred. No. 1.2e-19;

0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 TGCCTTGCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60

338 TGCCTTGCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 397

Qy

61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCTGTGT 101

398 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCTGTGT 438

Db

RESULT 2

AX952091

LOCUS

AX952091

DEFINITION

Sequence 1 from Patent WO03093485.

ACCESSION

AX952091

```

VERSION      AX952091.1  GI:40782473
KEYWORDS
SOURCE       synthetic construct
ORGANISM     other sequences; artificial sequences.
REFERENCE    1
AUTHORS      Mallet,J., Brun,S., Dufour,N. and Faucon-Biguet,N.
TITLE        Optimization of transgene expression in mammalian cells
JOURNAL      Patent: WO 03093485-A 1 13-NOV-2003;
              Centre National De La Recherche Scientifique-CNRS (FR) ; Biovectys
              (FR)
FEATURES     source
              1..609
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Description of Artificial Sequence: Primer"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGGCCCTGCTGTGACAGAGGCGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60
Db 338 TGCCTTGGCCCTGCTGTGACAGAGGCGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 397
Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGTCGCCCTGTGT 101
Db 398 TCGGGGAAGCTGACGCTCTTTCCATGGCTGTCGCCCTGTGT 438

RESULT 3
OHVSAGA
LOCUS
DEFINITION    OHVSAGA 1973 bp mRNA linear VRL 06-JUL-1995
              cds, polypeptide, 3' end, and cor protein, 5' end.
ACCESSION    M15954
VERSION      M15954.1  GI:893289
KEYWORDS     core protein; polymerase; surface antigen.
SOURCE       Woodchuck hepatitis B virus
ORGANISM     Woodchuck hepatitis B virus
REFERENCE    1 (bases 1 to 1973)
AUTHORS      Etiennele,J., Moroy,T., Trepo,C., Tiollais,P. and Buendia,M.A.
TITLE        Nucleotide sequence of the woodchuck hepatitis virus surface
              antigen mRNAs and the variability of three overlapping viral genes
              Gene 500, 207-214 (1986)
JOURNAL      87219879
MEDLINE
COMMENT      On Jul 7, 1995 this sequence version replaced gi:336154.
              Original source text: Woodchuck hepatitis B virus (strain W64)
              (clone: pWS23) cDNA to mRNA.
              Draft entry and printed copy of sequence [1] kindly provided by
              J.Etiennele, 22-JUN-1987.
FEATURES     source
              1..1973
               /organism="Woodchuck hepatitis B virus"
               /mol_type="mRNA"
               /strain="W64"
               /specific_host="Marmota monax liver"
               /db_xref="taxon:35269"
               /clone="pWS23"
               <1..1673
               /codon_start=3
               /product="polymerase"
               /protein_id="AAA69573.1"
               /db_xref="GI:336155"
               /translation="SKRYSPPLNVEKSDPSGVRGRIIRLDNNGTTPQCLWRSFYNT
               KPCGSCYCHHIVSSLDWGPCTVGDVITKSPRPRTITGGVFLVDKPNNSSESLRV
               LVDSQFSGRHTRWHPKFAVNLQTLANLLSTNLQWLSDVSAFYHIPISPAAVPHL
               LVGSPGLERFTTCLSSSTHNGDSQLOTHALCTRHVYSSLLILFKTYGRKLHLAHP
               FIMGFKLPMGVLSPLLAQFTSALASVMRRNPFHCVFAYMDLVLGARTSEHLTA
               IYSHICSVFLDLGITHLVNVTKNWGNHLMFGVYITSSGVLPPQDKHKVKLRYLRSVP
               "
CDS
              1..1973
               /product="surface antigen"
              55..1973
               /note="S mRNA (alt.)"
              63..1973
               /note="S mRNA (alt.)"
              72..1973
               /note="S mRNA (alt.)"
              82..1973
               /note="S mRNA (alt.)"
              1418..1843
               /note="putative"
               /codon_start=1
               /product="X protein"
               /protein_id="AAA69575.1"
               /db_xref="GI:893290"
               /translation="MAARLCCQLDSARDVLLRLPIQSSGPPFPFPAAGSAASSASS
               PPSDESDDLPLGLPACFASAGPCLVFTCADLRTMDSTVNFVSHAKRQLGMPKSD
               LMTPIKQDLTLKWEESIDPLRSIFVLGCRHKWRL"
              1936..1973
               /note="precursor"
               /codon_start=1
               /product="core protein"
               /protein_id="AA69576.1"
               /db_xref="GI:555278"
               /translation="MDIDPYKEFGSS"
               <1973..1973
               /product="WHV mRNA"
              ORIGIN      Unreported.
              Query Match      100.0%; Score 101; DB 14; Length 1973;
              Best Local Similarity 100.0%; Pred. No. 1.2e-19;
              Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              Qy 1 TGCCTTGGCCCTGCTGTGACAGAGGCGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60
              Db 1335 TGCCTTGGCCCTGCTGTGACAGAGGCGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 1394
              Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGTCGCCCTGTGT 101
              Db 1395 TCGGGGAAGCTGACGCTCTTTCCATGGCTGTCGCCCTGTGT 1435

RESULT 4
AF410858
LOCUS
DEFINITION    Woodchuck hepatitis B virus isolate 342 defective polymerase gene,
              complete cds.
ACCESSION    AF410858
VERSION      AF410858.1  GI:15637592
KEYWORDS     Woodchuck hepatitis B virus
SOURCE       Woodchuck hepatitis B virus
ORGANISM     Woodchuck hepatitis B virus
              2649 bp DNA linear VRL 18-FEB-2004
              AF410858
              Woodchuck hepatitis B virus isolate 342 defective polymerase gene,
              complete cds.
              AF410858
              Woodchuck hepatitis B virus
              GI:15637592
              Woodchuck hepatitis B virus
              Woodchuck hepatitis B virus
              ORGANISM
```

Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 2649)  
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.

TITLE Mutations of the woodchuck hepatitis virus polymerase gene that confer resistance to lamivudine and

JOURNAL 2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil

MEDLINE J. Virol. 76 (3), 1213-1223 (2002)

PUBMED 21635500

REFERENCE 11773397

AUTHORS 2 (bases 1 to 2649)

Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.

Direct Submission

Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme

Avenue, Philadelphia, PA 19111, USA

Location/Qualifiers

FEATURES 1. .2649

/organism="Woodchuck hepatitis B virus"

/virion

/mol\_type="genomic DNA"

/isolate="342"

/db\_xref="taxon:35269"

/note="isolated from serum of chronically infected

woodchuck"

1. .2649

/note="reverse transcriptase; contains six nucleotide

inframe deletion"

/codon\_start=1

/product="defective polymerase"

/protein\_id="AA04546.1"

/db\_xref="GI:15637593"

/translation="MHFPSRLFRNIQSLGEEVQELGPPEDALPLLAGEDLNHRVAD

ALNHLPTADLQWVHTNAITGLYSNOAOFNPHWIOPEFPELHLHNDLIQKQOYFG

PLTINERKQLQNPAPFEPKATKYFPLIKGINKNYENFALEHFFATANYLWTLWEAG

ILYLKRNQTLTLFKGPKYSWEHRLQVHQHQHSHLQSRQNSMWACSGYLLHNHLP

SEPVSVTRNLNNISDKSQRSTRTGLCSYKQIQTDRLHLARISCSGKITIGQGGSS

PKTSSISNFRNQTWAYNSRNSGHTTWFSSASNSKRSREKAYSSNSTSKRYSP

NYKSDSPSPGVRGRIITRLDNNGTTPQCLWRSFYNTKPGCGSYCTHHIVSLDDWGPCT

VTGDTVTIKSPTRPRITGGVFLVDKNPNNSSESLVDFQSFGRGHTRVHWPFAVEN

LOTLANLLSTNLQWLDVSAAFVHIPISTAVPHLLVGSFGLRFTTCLSSSTHNG

DSQLOQWHTLCTRHYSLLLLFTYGRKHLHLAHPFMGRKLPNGVGLSPFLLAQF

TSALASWVRNPHCVVFAYMDLDVLGARTSEHLTAIYSHICSVFLDLGLHNLNVRK

WMGNLHFMGVYITSSGVLQDQKHVKLSRYLSRVPNQPLDYKICERLTGLILNYAP

FTLCGYAALMPLVHAIASRTAFIFSSLYKSWLLSYBELWVVRQGVCTVADATP

TGCGIATTCOLLSTGTFAPLPATAEILAAACLRCTGARLLGTGDNVSLGKLTSP

FWLLACVANWILRGTSFCVPSALNPADLPSPGLLPVLRLPLRLRPQTSRLCLWA

PPVSPFRVRVAVSSPVQCEPWIPP"

Query Match 100.0%; Score 101; DB 14; Length 2649;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60

Db 2311 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 2370

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 101

Db 2371 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 2411

RESULT 5

AF410857

LOCUS

DEFINITION Woodchuck hepatitis B virus isolate 335 type I mutant polymerase

gene, complete cds.

ACCESSION AF410857

VERSION AF410857.1 GI:15637590

KEYWORDS

SOURCE Woodchuck hepatitis B virus

ORGANISM Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.

1 (bases 1 to 2655)

Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.

Mutations of the woodchuck hepatitis virus polymerase gene that confer resistance to lamivudine and

2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil

J. Virol. 76 (3), 1213-1223 (2002)

21635500

11773397

2 (bases 1 to 2655)

Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.

Direct Submission

Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme

Avenue, Philadelphia, PA 19111, USA

Location/Qualifiers

FEATURES 1. .2655

/organism="Woodchuck hepatitis B virus"

/virion

/mol\_type="genomic DNA"

/isolate="335"

/db\_xref="taxon:35269"

/note="lamivudine resistant

isolated from serum of chronically infected woodchuck"

1. .2655

/note="reverse transcriptase"

/codon\_start=1

/product="type I mutant polymerase"

/protein\_id="AA04545.1"

/db\_xref="GI:15637591"

/translation="MHFPSRLFRNIQSLGEEVQELGPPEDALPLLAGEDLNHRVAD

ALNHLPTADLQWVHTNAITGLYSNOAOFNPHWIOPEFPELHLHNDLIQKQOYFG

PLTINERKQLQNPAPFEPKATKYFPLIKGINKNYENFALEHFFATANYLWTLWEAG

ILYLKRNQTLTLFKGPKYSWEHRLQVHQHQHSHLQSRQNSMWACSGYLLHNHLP

SEPVSVTRNLNNISDKSQRSTRTGLCSYKQIQTDRLHLARISCSGKITIGQGGSS

PKTSSISNFRNQTWAYNSRNSGHTTWFSSASNSKRSREKAYSSNSTSKRYSP

NYKSDSPSPGVRGRIITRLDNNGTTPQCLWRSFYNTKPGCGSYCTHHIVSLDDWGP

VTGDTVTIKSPTRPRITGGVFLVDKNPNNSSESLVDFQSFGRGHTRVHWPFAVEN

LOTLANLLSTNLQWLDVSAAFVHIPISTAVPHLLVGSFGLRFTTCLSSSTHNG

DSQLOQWHTLCTRHYSLLLLFTYGRKHLHLAHPFMGRKLPNGVGLSPFLLAQF

TSALASWVRNPHCVVFAYMDLDVLGARTSEHLTAIYSHICSVFLDLGLHNLNVRK

WMGNLHFMGVYITSSGVLQDQKHVKLSRYLSRVPNQPLDYKICERLTGLILNYAP

FTLCGYAALMPLVHAIASRTAFIFSSLYKSWLLSYBELWVVRQGVCTVADATP

TGCGIATTCOLLSTGTFAPLPATAEILAAACLRCTGARLLGTGDNVSLGKLTSP

FWLLACVANWILRGTSFCVPSALNPADLPSPGLLPVLRLPLRLRPQTSRLCLWA

ASPSPFRVRVAVSSPVQCEPWIPP"

Query Match 100.0%; Score 101; DB 14; Length 2655;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60

Db 2317 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 2376

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 101

Db 2377 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 2417

RESULT 6

AF410859

LOCUS

DEFINITION Woodchuck hepatitis B virus isolate 335 polymerase gene, complete

cds.

ACCESSION AF410859

VERSION AF410859.1 GI:15637594

KEYWORDS

SOURCE Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 2655;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60

Db 2317 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 2376

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 101

Db 2377 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 2417

RESULT 6

AF410859

LOCUS

DEFINITION Woodchuck hepatitis B virus isolate 335 polymerase gene, complete

cds.

ACCESSION AF410859

VERSION AF410859.1 GI:15637594

KEYWORDS

SOURCE Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 2655;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60

Db 2317 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 2376

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 101

Db 2377 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 2417

RESULT 6

AF410859

LOCUS

DEFINITION Woodchuck hepatitis B virus isolate 335 polymerase gene, complete

cds.

ACCESSION AF410859

VERSION AF410859.1 GI:15637594

KEYWORDS

SOURCE Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 2655;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60

Db 2317 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 2376

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 101

Db 2377 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCCTGTGT 2417

RESULT 6

AF410859

LOCUS

DEFINITION Woodchuck hepatitis B virus isolate 335 polymerase gene, complete

cds.

ACCESSION AF410859

VERSION AF410859.1 GI:15637594

KEYWORDS

SOURCE Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

Woodchuck hepatitis B virus

REFERENCE

AUTHORS

```

REFERENCE 1 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and
Mason,W.S.
TITLE Mutations of the woodchuck hepatitis virus polymerase gene that
confer resistance to lamivudine and
2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
J. Virol. 76 (3), 1213-1223 (2002)
JOURNAL MEDLINE 21635500
PUBMED 11773397
REFERENCE 2 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and
Mason,W.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme
Avenue, Philadelphia, PA 19111, USA
FEATURES
source
1. .2655
/organism="Woodchuck hepatitis B virus"
/virion
/mol_type="genomic DNA"
/isolate="335"
/db_xref="taxon:35269"
/note="isolated from serum of chronically infected
woodchuck"
CDS
1. .2655
/note="reverse transcriptase"
/codon_start=1
/product="polymerase"
/protein_id="AAL04547.1"
/db_xref="GI:15637595"
/translaton="MHPFSRLFRNIQSLGREGVQELLGPPEDALPLIAGEDLNHRVAD
ALNHLPTADLQVHWKTNALTGLYSNOAQPNPHWIOPEPELHLDLQKLOQYFG
PLTINERKQLQNFPAFPFKATYFPLIKIGNNTPFALEHFFATANYLWTLWEAG
ILYLRKNQTTLTFKGPYSWEHRQLVQHQKHSHLQSQNSMVCACGYLLHNLHP
SEPVSVTRNLNNISDKQSKSTRGLCSYKQVQTDRLHLARISCGSKITIGQGGSS
PKTSYKSISSNFRNOTWAYNSRNSGHTWFFSASNSKRSREKAYSSNSTKRYSP
PLNYEKSDPSSPGVRGRIIRLDNNGTPQCLWBSFYNTKPGSYCIHHIVSSLDWGP
CTVGDVTIKSPRPRITGCVFLVDKNPNNSSESLVDFPSFGHTRVHWPKFAY
PNQTLANLSTNLQWLSLQVSAFYHIPISPAAPVHLLVGSGLERFVCLSSSTHN
GNDSQLQMTHTLCTRHLYSSLLFLKTYGRKLHLAHPFMGRKLPVGLSPFLLA
QFTSALASVMRNPFCVFPAYMDLVLGARTSEHLTAIYSHICSVPLDGLHNLNVK
TKWGNHLHFMGVYITSSGVLPODKVKLSRYLRSVPVNOPLDYKICERLTGLNVK
APFTLCGYAALMPLYHAIASRTAFIESLYKSWLLSLEYELWPVRQGVCTVFEADA
TPTGNGIATTCQLLSTGTFAPFLPFIATABLIAACLRACWTGARLLGTDNSVVLGKLTIS
FPMULLACVANWILRGTSFCYVPSALNPADLPVLRGLLPVLRPLRLRPLRPTQTSRLSMA
ASPPVSPRRPRVRAVWSSPVQTCPEWIPP"
ORIGIN
Query Match 100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGCCTGCTGACAGAGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60
Db 2317 TGCCTTGCCTGCTGCTGACAGAGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 2376
Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGCTGCTCCCTGTGT 101
Db 2377 TCGGGGAAGCTGACGCTCTTTCCATGCTGCTCCCTGTGT 2417
RESULT 7
AF410860
LOCUS Woodchuck hepatitis B virus isolate 335 clone b polymerase gene,
DEFINITION complete cds.
ACCESSION AF410860
VERSION AF410860.1 GI:15637596
KEYWORDS Woodchuck hepatitis B virus
SOURCE Woodchuck hepatitis B virus
ORGANISM Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 2655)

```

```

AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and
Mason,W.S.
TITLE Mutations of the woodchuck hepatitis virus polymerase gene that
confer resistance to lamivudine and
2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
J. Virol. 76 (3), 1213-1223 (2002)
JOURNAL MEDLINE 21635500
PUBMED 11773397
REFERENCE 2 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and
Mason,W.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme
Avenue, Philadelphia, PA 19111, USA
FEATURES
source
1. .2655
/organism="Woodchuck hepatitis B virus"
/virion
/mol_type="genomic DNA"
/isolate="335"
/db_xref="taxon:35269"
/note="isolated from serum of chronically infected
woodchuck"
CDS
1. .2655
/note="reverse transcriptase"
/codon_start=1
/product="polymerase"
/protein_id="AAL04548.1"
/db_xref="GI:15637597"
/translaton="MHPFSRLFRNIQSLGREGVQELLGPPEDALPLIAGEDLNHRVAD
ALNHLPTADLQVHWKTNALTGLYSNOAQPNPHWIOPEPELHLDLQKLOQYFG
PLTINERKQLQNFPAFPFKATYFPLIKIGNNTPFALEHFFATANYLWTLWEAG
ILYLRKNQTTLTFKGPYSWEHRQLVQHQKHSHLQSQNSMVCACGYLLHNLHP
SEPVSVTRNLNNISDKQSKSTRGLCSYKQVQTDRLHLARISCGSKITIGQGGSS
PKTSYKSISSNFRNOTWAYNSRNSGHTWFFSASNSKRSREKAYSSNSTKRYSP
PLNYEKSDPSSPGVRGRIIRLDNNGTPQCLWBSFYNTKPGSYCIHHIVSSLDWGP
CTVGDVTIKSPRPRITGCVFLVDKNPNNSSESLVDFPSFGHTRVHWPKFAY
PNQTLANLSTNLQWLSLQVSAFYHIPISPAAPVHLLVGSGLERFVCLSSSTHN
GNDSQLQMTHTLCTRHLYSSLLFLKTYGRKLHLAHPFMGRKLPVGLSPFLLA
QFTSALASVMRNPFCVFPAYMDLVLGARTSEHLTAIYSHICSVPLDGLHNLNVK
TKWGNHLHFMGVYITSSGVLPODKVKLSRYLRSVPVNOPLDYKICERLTGLNVK
APFTLCGYAALMPLYHAIASRTAFIESLYKSWLLSLEYELWPVRQGVCTVFEADA
TPTGNGIATTCQLLSTGTFAPFLPFIATABLIAACLRACWTGARLLGTDNSVVLGKLTIS
FPMULLACVANWILRGTSFCYVPSALNPADLPVLRGLLPVLRPLRLRPLRPTQTSRLSMA
ASPPVSPRRPRVRAVWSSPVQTCPEWIPP"
ORIGIN
Query Match 100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGCCTGCTGCTGACAGAGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60
Db 2317 TGCCTTGCCTGCTGCTGACAGAGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 2376
Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGCTGCTCCCTGTGT 101
Db 2377 TCGGGGAAGCTGACGCTCTTTCCATGCTGCTCCCTGTGT 2417
RESULT 8
AF410861
LOCUS Woodchuck hepatitis B virus isolate 342 polymerase gene, complete
DEFINITION cds.
ACCESSION AF410861
VERSION AF410861.1 GI:15637598
KEYWORDS Woodchuck hepatitis B virus
SOURCE Woodchuck hepatitis B virus
ORGANISM Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 2655)

```

AUTHORS	Yamanoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE	Mutations of the woodchuck hepatitis virus polymerase gene that confer resistance to lamivudine and 2'-fluoro-5-methyl-beta-L-arabino-furancosyluracil
JOURNAL	J. Virol. 76 (3), 1213-1223 (2002)
MEDLINE	21635500
PUBMED	11773397
REFERENCE	2 (bases 1 to 2655)
AUTHORS	Yamanoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE	Direct Submission
JOURNAL	Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA
FEATURES	Location/Qualifiers source 1..2655 /organism="Woodchuck hepatitis B virus" /viral /mol_type="genomic DNA" /isolate="342" /db_xref="taxon:35269" /note="isolated from serum of chronically infected woodchuck"
CDS	1..2655 /note="reverse transcriptase" /codon_start=1 /product="polymerase" /protein_id="AAL04549.1" /db_xref="GI:15637599" /translation="MPPFRSLFNNISLGEEVEVQELGPPEDALPLLAGEDLNMHRVAD ALNHLLTADLOQLWHPKNAITGLYSNOAAQFNPHIOPFEPELHENDLIQKLQQYFG PLTNIEKRKLQNFAPFPKATYFKLKGINKNYENFALEHFFATANYLWTLMWAG ILYLRKNQTTLTFKPXPYSMEHRQLVQHVGQOQKHLSQRSSMWACSGYLHNHLSP SEPVSYSTNLNNSNISDKSQSRVLGTCSVKYOIQDRLEHARI SCGSKITTGQOGSS PRTSYKSI SSNPERNOTWAYNSRSNGHTTWFSASNSKRSREKAYSNSNSTSKRYSP PLNYESKDRESSPCVRGRIIRLDNNGPTPQCLMRSFNTPKGSCYCIHHIVSSLDMDGP CIVTGDIKSPTRPRITGGVELVDKNPNNSSESLVDFQSFGHTRVHWMPKFAV PNLTQIANLSTLNWLSDVSAFYHIPISPAAPHLVLVSGPSLERFTTCISSSTHAN GNDSQOQWHITLCTRHVYSLLILFKTYGRKLHLHAHPFMGPRKPLMGVGLSPFLLA QTSALASVMRRNPFCVSVAYMDMLVLGARTSEHLTAYSHICSVFLDLGHANVNK TQWMGNLFHMFGVVITSSGVLPQDHVKLSRVLSVPVNPQDLYKICERLGTILNVY APTLCGYAALMPLHYAIASRTAFISLYSWILLSBELMPVVRQORVCVTGFADA TPTWGIAITCOLLSGTAPPPIAPATAELIAACLARCWTGARLLGDTPDNSVLSGKITS FWLLACVANWILRGTSFCVPSALNPADLPISRGLLPVLRPLRLRLRQPOTSRISLWA ASPPVSPRPAPARVAWSSPYQCETCPWIPP"
ORIGIN	
Query Match	100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity	100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TGCCCTTGCCCGCTGCTGGACAGGGCTTCGCCTTTGGGCACATCAATTCCTGGTGTG 60
Dd	2317 TGCCCTTGCCCGCTGCTGGACAGGGCTTCGCCTTTGGGCACATCAATTCCTGGTGTG 2317
QY	61 TCGGGGAAGCTGACGCTCCTTTCCATGCCTGCTCGCCTGTGT 101
Dd	2377 TCGGGGAAGCTGACGCTCCTTTCCATGCCTGCTCGCCTGTGT 2417
RESULT 9	
OHVCGB	3320 bp ms-DNA circular VRL 22-JUN-1994
LOCUS	Woodchuck hepatitis virus 2, complete genome.
DEFINITION	ML1082
ACCESSION	ML1082.1 GI:336132
VERSION	DNA polymerase; Wbc protein; WbsAg protein; coat protein; complete genome; core protein; env protein; origin of replication; polymerase
KEYWORDS	Woodchuck hepatitis B virus
SOURCE	Woodchuck hepatitis B virus
ORGANISM	Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE	1 (bases 1 to 3320)

/translation="MDIDPYKEFGSSVQLNPLDFFPDNLALVDTATLYEELTG  
REHCSPHHTAIRQALVCMDELTKLIAMSSNITSEQVRIIVNHVNDTWGLKYRQSLM  
FHLSCUTFGHTVQBEFLVSFWIRTPAPTRPNAPILSTLPEHTVIRRGARASRP  
RRRTSPRRRSQSPRRRSQSPSANC"  
ORIGIN EcoRI site.

Query Match 100.0%; Score 101; DB 14; Length 3320;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 60  
Db 1420 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 1479  
Qy 61 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 101  
Db 1480 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 1520

RESULT 10  
OHVCGD 3323 bp ms-DNA circular VRL 04-MAY-1994  
LOCUS Woodchuck hepatitis virus (WHV), complete genome, clone WHV 59.  
DEFINITION M19183  
ACCESSION M19183  
VERSION M19183.1 GI:336141  
KEYWORDS DNA polymerase; WHC protein; WHAag protein; coat protein; complete genome; core protein; envelope-associated protein; origin of replication; polymerase.  
SOURCE Woodchuck hepatitis B virus  
ORGANISM Woodchuck hepatitis B virus  
VIRUSES: Retroviruses; Hepadnaviridae; Orthohepadnavirus.  
REFERENCE 1 (bases 1 to 3323)  
AUTHORS Cohen, J.I., Miller, R.H., Rosenblum, B., Denniston, K., Gerin, J.L. and Purcell, R.H.  
TITLE Sequence comparison of woodchuck hepatitis virus replicative forms  
JOURNAL shows conservation of the genome  
MEDLINE Virology 162 (1), 12-20 (1988)  
PUBMED 88101359  
COMMENT Original source text: Woodchuck hepatitis virus DNA, clone WHV 59, from a carrier woodchuck trapped in Pennsylvania.  
FEATURES  
source  
1..3323  
/organism="Woodchuck hepatitis B virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35269"  
join(2427..3323,1..1758)  
/codon\_start=1  
/product="DNA polymerase"  
/protein\_id="AAA46763.1"  
/db\_xref="GI:336143"  
/translation="MHPFSRLFRNIQSLGEEVQELGPPEDALPLLAGEDLNHRVAD  
ALNHLPTADLQWHTKNAITGLYSNQAQFNPHWIOPEFELHNDLTKLQYVFG  
PLTNEKRLQLNPAPFPFKATYVFLIKGINKNYPNFALEHFFATANYLWTLWAG  
ILVLRKQTLTKFKGYSMEHRLQVHNGQKHLSQSRNSWACSGHLLNHLPL  
SEPVSTRNLNSIDKSQKTKGLCSYKQVTDREHARISCSKSTIQCGSS  
PKTSYKISNFRNQTWAYNSNSGHTWFSASNSKRSREKAYSNSITSORSP  
CLTVGTDTIKSPTRPTITGLVDKNPNNSSESLVDFVSQFSRGRHVRHWPKEAV  
PNLQTLANLNLQSLDVSAAFYHIPISPAAPHLLVSGPLERFNTCMSSSTHN  
GNDSLOLTHALCTRHVYSSILLFLFKTYGRKLHLAHFFIMGRKLPNGVGLSPFLIA  
QFTSAISMVARNRPFHCVFAYMDLVGARTSEHLTAIYSHICSVFLDGLHNVNK  
TKWGNHLHFMGYVITSSGLPQDKHVKLSRYLRSVPNQPDLQDKICERLTGLNVY  
APFTLCGYALPHYIAISRTAFIFSSLYKSWLLSYELWLPVVRQGVCTVFADA  
PTTCGWTATTCQLLSGTFAPPLPIATAELIAACIARLCTGRLIGTNSVVLGSKLTS  
PFWLLACVANNILGTSFCVPSALNPADLPSCRLLPVLRPLRLRLPQTSRLSLWA  
ASPVSPRRPVRAWSPVQCEPWIP"  
join(2992..3323,1..964)  
/note="precursor"  
/codon\_start=1  
/product="surface protein"  
/protein\_id="AAA46762.1"  
/db\_xref="GI:336142"

CDS

/translation="MGNNIKVTFNPKIAAWPAVGTYYTITVTPQNSVFPQGIYOTT  
SLINPKNOELDSVLINRYKQIDMNTWQGFVVDOKLPLVSRDPLAPHLNOSAOETEI  
KFGPIVIFGRIDIPRGLVFPQPTNRDQGRKPTPTPLRDTHPHLTMKNOTRLOGF  
VGLRDLITTERIHNAIDPFITLSPVFTSTLSPSTITGDPALSPSPSSLLGL  
LAGLQVYFLMTIKLITIAQNLDWMWTSLSFPGGIPECTGQNSQFQTKHLPTSCPTC  
NGFRWMLRRFIIYLLVLLCLIFLLVLDWKLIFVCPLOPTTETTVNCRQCTLSVQ  
DVTVPYCCCLKPTAGNCTCWPDPSPSWALGNLYLMEWALARFSLNLLVPLLOWLGGIS  
LTAWELLIMWFMGPAALLSILPPRIPIFVLFLIWWYI"  
join(2992..3323,1..295)  
/notes="surface protein"  
296..961  
/product="surface protein"  
1503..1928  
/codon\_start=1  
/product="X protein"  
/protein\_id="AAA46764.1"  
/db\_xref="GI:336144"  
/translation="NAARLCCQLDSARDVLLRPFQPSQSGPPFPRPAAGSAASSTSS  
PSPDESULPLGLRIPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKRQLGMPKSD  
LWTPYIKDOLLTKWEEGSDPRLSIFVLGGRHKWRLL"  
2021..2587  
/codon\_start=1  
/product="core protein"  
/protein\_id="AAA46765.1"  
/db\_xref="GI:336145"  
/translation="MDIDPYKEFGSSVQLNPLDFFPDNLALVDTATLYEELTG  
REHCSPHHTAIRQALVCMDELTKLIAMSSNITSEQVRIIVNHVNDTWGLKYRQSLM  
FHLSCUTFGHTVQBEFLVSFWIRTPAPTRPNAPILSTLPEHTVIRRGARASRP  
RRRTSPRRRSQSPRRRSQSPSANC"  
ORIGIN  
Query Match 100.0%; Score 101; DB 14; Length 3323;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 60  
Db 1420 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 1479  
Qy 61 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 101  
Db 1480 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 1520  
RESULT 11  
OHVCGD 3323 bp ms-DNA circular VRL 04-MAY-1994  
LOCUS Woodchuck hepatitis virus (WHV), complete genome, clone WHV 7.  
DEFINITION M18752  
ACCESSION M18752  
VERSION M18752.1 GI:336136  
KEYWORDS DNA polymerase; coat protein; complete genome; core protein; envelope-associated protein; origin of replication; polymerase.  
SOURCE Woodchuck hepatitis B virus  
ORGANISM Woodchuck hepatitis B virus  
VIRUSES: Retroviruses; Hepadnaviridae; Orthohepadnavirus.  
REFERENCE 1 (bases 1 to 3323)  
AUTHORS Cohen, J.I., Miller, R.H., Rosenblum, B., Denniston, K., Gerin, J.L. and Purcell, R.H.  
TITLE Sequence comparison of woodchuck hepatitis virus replicative forms  
JOURNAL shows conservation of the genome  
MEDLINE Virology 162 (1), 12-20 (1988)  
PUBMED 88101359  
COMMENT Original source text: Woodchuck hepatitis virus DNA, clone WHV 7, from a carrier woodchuck trapped in Maryland.  
FEATURES  
source  
1..3323  
/organism="Woodchuck hepatitis B virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35269"  
join(2427..3323,1..1758)  
/codon\_start=1  
/product="DNA polymerase"

/protein\_id="AAA46767.1"  
/db\_xref="GI:336138"  
/translation="MHFSLRFRNIQSLGEEVEQELGPPEDALPLLAGEDLNHRVAD  
ALNHLPTADLOWHKNTAITGLYSNOAQFNPHWIOPEFELHLHNDLLOKLOQYFG  
PLTNEKRLQDLPNPAFFPKATKYFPLIKGINKNYENFALHFFATANYLWTLWEAG  
ILYLRKQNTLTFGKQYSWEHROLVHNGOQSHLQSRQNSWVACSHLHNLHS  
SESVSVTRNLNNISDKSQSTRIGLCSYKQIOTDRLEHRLARISCSKGTIGQGGSS  
PKTIJYSSFNQWYNSSRNSGHTWFFSSNSKSRREKAYSSNSTSKRYP  
PLNYEKDFSPGVRITRDLNNGTPTQCLMRSFYNTKPGCYICHIIVSSLDWGP  
CTVVDVTKISPRPTITGGVFLVDKNPNNSSRVLPVDFQSFGRHTRVHWPKFV  
PNCOTLANLLSLQWLDVSAFYHPIPIPAAPVHLLVSGPLERFNTCLSSSTN  
RNSQLOTMNLCTRHVYSILLFKTVGRKLHLLAHFIMGFELPMGVCLSPFLA  
OPTSALSMVRNRPCHVCFVAYMDLVLGARTSEHLTAIYSHICSVFELDLGHLNVK  
TKWGNHLFMFYVITSSGFLPQDKVKKISRYLRSVFNQPLDYKICERITGLTIV  
APFTLCGYAALMPLYHATISRTAFISLSYKSWLLSYBELVPPVRQGVCTVAFDA  
TPTGGLATTYLLGFAPFLPIATAELIAACIARCTGARTLGTDNSVLSGKLT  
FWPWLACVANWILGTFQVPSALNPADLPGLPVLRPLRLRLRPQTSRISLWA  
ASPVSPRRPVVAVSSPVQCEWIPP"  
join(2992..3323,1..964)  
/note="precursor"  
/codon\_start=1  
/product="surface protein"  
/protein\_id="AAA46766.1"  
/db\_xref="GI:336137"  
/translation="MGNNIKVTNPDKIAWPAVGVYTYTTPQNSVFPQGIYQTT  
SLINKNQQLDVLINRYKQIDNTWQGVDPQKJPLVSRDPPKPYINQSAQTFEI  
KPGPIIVGRIIVDPGLVPPQPTNDRQGRKPTPTPLRTHPPLTKMQTFHLQGF  
VDGRLDTTTEROHNAVDPTTSLSPAVPTVSTLSPSTTGDPALESPSSLLQGF  
LAGIQQVYFLWTKILTIAQNLDMWMTLSLSPGGIPECTGONSOFQTKHLPTSCPPNC  
NGFRWYLRPIIYVLLVLLCLIFLLVLDWKGILPVCLOPTTETTVNCRQCTISAQ  
DMYTPPYCCCLTAGNCTWPISSWALGNLWELARFSLNLLVPLLLQWJGGIS  
LIAWFLLIWIMFWGPAALLSILPFIPIFVFLFIWYI"  
join(2992..3323,1..295)  
/note="pre-surface protein"  
296..961  
/product="surface protein"  
/codon\_start=1  
/product="X protein"  
/protein\_id="AAA46768.1"  
/db\_xref="GI:336139"  
/translation="MAARLCQLDSARDVLLRFPFGQSSGSPFPFPAAGSAASSASS  
PSPSDESLPLGLRPACFASAGPCCLVFTCAELRTWDSVFNFSWANHRLQGMPSKD  
LWTPYIKDQLLTKEEGSIDPRLSIFVLGGRHKMRLI"  
2021..2587  
/codon\_start=1  
/product="core protein"  
/protein\_id="AAA46769.1"  
/db\_xref="GI:336140"  
/translation="MDIDPYKFGSSYQLNLFPLDFFPDNLALVDATATAYEBELTG  
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW  
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS  
PRRTSPRRRSQSPRRRSQSPSANC"

## ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 3323;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 60  
Db 1420 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 1479  
Qy 61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCTGTGT 101  
Db 1480 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCTGTGT 1520

## RESULT 12

OHVHEPBA OHVHEPBA 3323 bp ss-DNA linear VRL 03-AUG-1993  
LOCUS Woodchuck hepatitis B virus (WHV8), complete genome.  
DEFINITION J04514  
ACCESSION

VERSION  
KEYWORDS

J04514.1 GI:336146  
Woodchuck hepatitis B virus  
Woodchuck hepatitis B virus

ORGANISM  
Viruses: Retroviridae; Hepadnaviridae; Orthohepadnavirus.

REFERENCE  
1 (bases 1 to 3323)

AUTHORS  
Girones,R., Cole,P.J., Hornbuckle,W.E., Tennant,B.C., Gerin,J.L.,  
Purcell,R.H. and Miller,R.H.

TITLE  
Complete nucleotide sequence of a molecular clone of woodchuck  
hepatitis virus that is infectious in the natural host

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1846-1849 (1989)

MEDLINE  
89184524

PubMed  
2328306

COMMENT  
Original source text: Woodchuck hepatitis B virus (strain WHV8),  
cDNA to viral DNA, clone pWHV8.

Draft entry and computer-readable sequence for [1] kindly submitted  
by R.Girones, 24-MAY-1989.

FEATURES  
source  
1..3323  
/organism="Woodchuck hepatitis B virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35269"  
296..964  
/note="surface protein"  
/codon\_start=1  
/protein\_id="AAA46770.1"  
/db\_xref="GI:336147"  
/translation="MSPSSLGLLAGLVVYVFLTKILTIAQNLDMWMTLSLSPGCGIP  
ECTGNSQFOTCKHLPTSCPTCNGFRMYLRRLIYVLLVLLCLIFLLVLDWKGILI  
PVCLOPTTETTVNCRQCTISADMTYPPCCCLKPTAGNCTWPISSWALGNLYWE  
WAFRWSWMLLVPLVQLWGLGILSIAMFLLIWMFWGPAALLSILPFIPIFVFLFI  
WYI"  
1503..1928  
/note="X protein"  
/codon\_start=1  
/protein\_id="AAA46771.1"  
/db\_xref="GI:336148"  
/translation="MAARLCCLDSARDVLLRFPFGQSSGSPFPFPAAGSAASSASS  
PSPSDESLPLGLRPACFASAGPCCLVFTCADLRTWDSVFNFSWANHRLQGMPSKD  
LWTPYIKDQLLTKEEGSIDPRLSIFVLGGRHKMRLI"  
1719..1728  
/note="direct repeat"  
1941..1951  
/note="direct repeat"  
2021..2587  
/note="core protein"  
/codon\_start=1  
/protein\_id="AAA46772.1"  
/db\_xref="GI:336149"  
/translation="MDIDPYKFGSSYQLNLFPLDFFPDNLALVDATATAYEBELTG  
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW  
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS  
PRRTSPRRRSQSPRRRSQSPSANC"

CDS  
1719..1728  
/note="direct repeat"  
1941..1951  
/note="direct repeat"  
2021..2587  
/note="core protein"  
/codon\_start=1  
/protein\_id="AAA46772.1"  
/db\_xref="GI:336149"  
/translation="MDIDPYKFGSSYQLNLFPLDFFPDNLALVDATATAYEBELTG  
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW  
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS  
PRRTSPRRRSQSPRRRSQSPSANC"

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

repeat\_region  
repeat\_region  
CDS

VERSION AY468486.1 GI:38565537  
KEYWORDS  
SOURCE Lenticlinal transfer vector pHSXW  
ORGANISM Lenticlinal transfer vector pHSXW  
REFERENCE 1 (bases 1 to 7207)  
AUTHORS Johansen,J., Tornoe,J., Rosenblad,C., Dago,L. and Kusk,P.  
TITLE Improved lenticlinal transfer vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 7207)  
AUTHORS Kusk,P. and Johansen,J.  
TITLE Direct Submission  
JOURNAL Submitted (19-NOV-2003) NsGene A/S, Baltorpvej 154, Ballerup DK-2750, Denmark

FEATURES  
source  
1..7207  
/organism="Lenticlinal transfer vector pHSXW"  
/mol\_type="other DNA"  
/db\_xref="taxon:256321"  
LTR  
1..634  
/note="5' HIV-1 LTR"  
misc\_signal  
686..823  
/note="packaging signal; psi"  
misc\_signal  
1310..1514  
/note="Rev responsive element; RRE"  
promoter  
2031..2547  
/note="CMV promoter"  
misc\_feature  
2618..2671  
/note="multiple cloning site"  
misc\_feature  
complement(2674..2693)  
/note="T7 recognition/binding site"  
misc\_feature  
2711..3315  
/note="Woodchuck post-regulatory element; WPRE"  
LTR  
3512..3746  
/note="3' LTR; self-inactivating 3' LTR version; distal promoter of U3 deleted"  
rep\_origin  
4719..5392  
complement(5537..6397)  
gene  
/gene="bla"  
CDS  
complement(5537..6397)  
/gene="bla"  
/note="ampicillin resistance gene"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="AA24091.1"  
/db\_xref="GI:38565538"  
/translation="MSIQHFRVALIPFPAFLPVFAHPETLVKVKDAEDQLGARVGY  
IEDLNSGKILESFRPRPEPMSTFKVLLCGAVLSRI DAGQQLGR LHYSONDLVE  
YSPTEKHLTDGTVRELCSAALTMSDNTAANLLITYGGPKELTAPFLHMGDHVTRL  
DRWEPELNEALPNDERTTPVAVATTIRKLTLTGELLTLASRQOLIDWMEADKVGPL  
LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVITYTGSQATMDERNRQIA  
EIGASLIKHW"

ORIGIN  
source  
Query Match 100.0%; Score 101; DB 12; Length 7207;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60  
Db 3048 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 3107  
Qy 61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCCTGTGT 101  
Db 3108 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCCTGTGT 3148

RESULT 14  
AX663053  
LOCUS 7515 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 4 from Patent WO02086132.  
ACCESSION AX663053

VERSION AX663053.1 GI:29163598  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Kingman,S.A., Mitrophanous,K. and Ellard,F.M.  
TITLE Vector system  
JOURNAL Patent: WO 02086132-A 4 31-OCT-2002;  
Oxford Biomedica (UK) Limited (GB)  
FEATURES  
source  
1..7515  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pSmart2 MCS 5prime cppt"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 7515;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60  
Db 4424 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 4483  
Qy 61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCCTGTGT 101  
Db 4484 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTCGCCTGTGT 4524

RESULT 15  
MAR7  
LOCUS 7808 bp DNA linear ROD 06-APR-1994  
DEFINITION Woodchuck hepatitis virus surface antigen (pres) and X protein genes.  
ACCESSION M60766  
VERSION M60766.1 GI:191471  
KEYWORDS X protein; surface protein.  
SOURCE Marmota monax (woodchuck)  
ORGANISM Marmota monax  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae; Marmota.  
REFERENCE 1 (bases 1 to 7808)  
AUTHORS Yamazoe,M., Nakai,S., Ogasawara,N. and Yoshikawa,H.  
TITLE Integration of woodchuck hepatitis virus (WHV) DNA at two chromosomal sites (Vk and gag-like) in a hepatocellular carcinoma Gene 100, 139-146 (1991)  
JOURNAL 91276235  
MEDLINE  
PUBMED 2055466  
COMMENT source text: Woodchuck hepatitis virus DNA, clone B7.  
FEATURES  
Original Location/Qualifiers  
source  
1..7808  
/organism="Marmota monax"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9995"  
/tissue\_type="liver"  
4123..4496  
/note="viral coding region"  
/codon\_start=3  
/product="surface antigen"  
/protein\_id="AAA37108.1"  
/db\_xref="GI:191472"  
/translation="GLLPVCPQLPTTETTVNCRQCTLSVQDITYTPPYCCCLKPTAGNC  
TCWPIPSWALGNLWELARPSWLNLLVPLQLWLGISLIANFLLIWMFWGPAL  
LSILPPIPIFVLFLLIWIYI"  
5035..5505  
/note="protein coding region crossing over viral/host boundary"  
/codon\_start=1  
/protein\_id="AAA37110.1"  
/db\_xref="GI:191474"



```

/translation="MAARLCCQLDSARDVLLRRFPQSSGPPPPRPAAGSAASSTSS
LSPSDESDLPGLRPLPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKQLGMPKSD
LWTFYIKDQLLTDQLLTIDNMQFYFKSPFLPSDFNVSTPALDPMGLSLQFWI"
5035..5392
/note="viral coding region"
/codon_start=1
/product="X protein"
/protein_id="AAA37109.1"
/db_xref="GI:553849"
/translation="MAARLCCQLDSARDVLLRRFPQSSGPPPPRPAAGSAASSTSS
LSPSDESDLPGLRPLPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKQLGMPKSD
LWTFYIKDQLLTDQLLT"
5590..5595
polyA_signal
```

ORIGIN

```

Query Match          100.0%; Score 101; DB 10; Length 7808;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGTCGTGTG 60
          |||
Db      4952 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGTCGTGTG 5011
          |||

Qy      61  TCGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101
          |||
Db      5012 TCGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 5052
          |||
```

Search completed: July 14, 2005, 14:03:37  
Job time : 757.618 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_4192\_4292  
Perfect score: 101  
Sequence: 1 tgccttcccgcgtgctgac.....ccatggctgctgcgtgtgt 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	592	9 ADA38373	Ada38373 Woodchuck
2	101	100.0	604	8 AAD55110	Aad55110 Woodchuck
3	101	100.0	604	8 ACC45094	Acc45094 Woodchuck
4	101	100.0	609	10 ADD67513	Add67513 WPRE elem
5	101	100.0	632	12 ADO25310	Ado25310 Woodchuck
6	101	100.0	2853	8 AAD55114	Aad55114 Hflapubig
7	101	100.0	2853	8 ACC45098	Acc45098 HIV-1 fla
8	101	100.0	7515	8 ABV77010	Abv77010 Nucleotid
9	101	100.0	7648	12 ADM47497	Adm47497 Lysosomal
10	101	100.0	8092	12 ADM47498	Adm47498 Lysosomal
11	101	100.0	8484	3 AAAS9091	Aaas9091 Nucleotid
12	101	100.0	8484	10 ADF48775	Adf48775 fibre exp
13	101	100.0	9731	11 ADM82791	Adm82791 DNA repai
14	101	100.0	9731	11 ADM82791	Adm82791 DNA repai
15	101	100.0	9782	11 ADM82792	Adm82792 DNA repai
16	101	100.0	10468	9 ACD27899	Acd27899 pdm12 vec
17	94.6	93.7	592	2 AAAX32299	Aax32299 Nucleotid
18	94.6	93.7	632	12 ADO01085	Ado01085 Woodchuck
19	94.6	93.7	6893	10 ADE24111	Ade24111 Proviral
20	91.4	90.5	3671	6 AAD28271	Aad28271 Alpha-lac

21	91.4	90.5	5617	6 AAD32077	Aad32077 Human alp
22	91.4	90.5	5691	6 AAD28313	Aad28313 Alpha-lac
23	91.4	90.5	5691	6 AAD28274	Aad28274 Alpha-lac
24	91.4	90.5	5691	12 ADM68974	Adm68974 Alpha-lac
25	91.4	90.5	5711	6 AAD28310	Aad28310 Alpha-lac
26	91.4	90.5	5711	12 ADM68971	Adm68971 Alpha-lac
27	91.4	90.5	5731	6 AAD32078	Aad32078 Human alp
28	91.4	90.5	5732	6 AAD28308	Aad28308 Mouse mam
29	91.4	90.5	5732	6 AAD28269	Aad28269 Mouse mam
30	91.4	90.5	5732	12 ADM68969	Adm68969 MMTV MN14
31	91.4	90.5	6026	6 AAD32075	Aad32075 Human alp
32	91.4	90.5	6027	12 ADL35206	Adl35206 Plasmid p
33	91.4	90.5	6140	6 AAD32076	Aad32076 Human alp
34	91.4	90.5	6706	12 ADL35208	Adl35208 Plasmid p
35	91.4	90.5	6748	12 ADL35207	Adl35207 Plasmid p
36	91.4	90.5	7033	12 ADO07393	Ado07393 Modified
37	91.4	90.5	7248	12 ADL35209	Adl35209 Plasmid p
38	91.4	90.5	7350	12 ADL35212	Adl35212 Plasmid p
39	91.4	90.5	7650	12 ADL35213	Adl35213 Plasmid p
40	91.4	90.5	7927	12 ADL35211	Adl35211 Plasmid p
41	91.4	90.5	7969	12 ADL35210	Adl35210 Plasmid p
42	91.4	90.5	9183	6 AAD28309	Aad28309 Alpha-lac
43	91.4	90.5	9183	6 AAD28270	Aad28270 Alpha-lac
44	91.4	90.5	9183	12 ADM68970	Adm68970 Alpha-lac
45	91.4	90.5	9291	10 ACH00963	Ach00963 PSMA anti

ALIGNMENTS

RESULT 1  
ADA38373  
ID - ADA38373 standard; DNA; 592 BP.  
XX ADA38373;  
AC ADA38373;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Woodchuck hepatitis virus postranscriptional regulatory element.  
XX  
KW woodchuck hepatitis virus; postranscriptional regulatory element; ds;  
KW trans-lenti viral vector system; Vpr; Vpx; reverse transcriptase;  
KW integrase; gag; gag-pol precursor;  
KW polypurine tract-central terminator sequence; PPT-CTS; gene transfer;  
KW gene expression; transduction; replication competent retrovirus; RCR.  
XX  
OS Woodchuck hepatitis B virus.  
XX  
FN US2003072938-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 24-JUL-2002; 2002US-00202457.  
XX  
PR 03-JUN-1998; 98US-00089900.  
PR 14-DEC-1999; 99US-00460548.  
PR 13-NOV-2000; 2000US-00709501.  
XX  
PA (KAPP/) KAPPES J C.  
PA (WUXX/) WU X.  
XX  
PI Kappes JC, Wu X;  
XX  
XX WPI; 2003-625514/59.  
XX  
PT New trans-lenti viral vector systems with cis-acting sequences.  
PT polypurine tract-central terminators, etc., useful for gene transfer,  
PT e.g. delivering genes encoding therapeutic or viral inhibitory  
XX polypeptides to cells.  
XX  
PS Example 5; Page 7; 19pp; English.  
XX  
CC The invention describes a trans-lenti viral vector system. The new trans-

CC lenti viral vector system comprises: a first nucleic acid segment  
CC encoding at least one fusion protein comprising a functional portion of a  
CC Vpr or Vpx polypeptide fused in frame to a functional portion of a  
CC Reverse transcriptase polypeptide fused in frame to a functional portion of a  
CC of an integrase polypeptide, where the first nucleic acid segment is  
CC capable of expression in a mammalian cell, and where the functional  
CC portion of the Vpr or Vpx polypeptide is capable of providing for the  
CC incorporation of the fusion protein into a viral particle; a second  
CC nucleic acid segment comprising a nucleotide sequence encoding a  
CC functional portion of a Gag polypeptide and functional portion of a  
CC Protease polypeptide, where the second nucleic acid is capable of  
CC expression in the mammalian, and where the second nucleic acid does not  
CC encode a functional Reverse Transcriptase polypeptide or a functional  
CC Integrase polypeptide; a third nucleic acid segment comprising a nucleic  
CC acid sequence encoding a functional envelope polypeptide capable of  
CC mediating recognition and entry of the viral particle into a target cell,  
CC where the third nucleic acid segment does not encode a functional Gag-Pol  
CC precursor; and a fourth nucleic acid segment comprising a heterologous  
CC nucleic acid sequence and at least one nucleotide sequence consisting of  
CC a functional equivalent of a polypurine tract-central terminator sequence  
CC (PPT-CTS), a functional equivalent of Woodchuck Hepatitis Virus Post  
CC transcriptional Regulatory Element (WPRE), a PPT-CTS or a WPRE. The trans  
CC -viral vector system produces a viral particle capable of introducing the  
CC heterologous nucleotide sequence into the genome of the target cell. The  
CC target cell is a non-dividing cell, a primary cell, macrophage, a CD34+  
CC cell, or haematopoietic stem cell. The trans-lenti viral vector system is  
CC useful for gene transfer, particularly for producing viruses or viral  
CC particles capable of introducing heterologous nucleotide sequence(s) into  
CC the genome of a target cell. The trans-lenti viral vector system is  
CC particularly useful for delivering heterologous nucleic acid sequences  
CC (which encode viral inhibitory polypeptides or a therapeutic  
CC polypeptides) to cells with greater efficiency and/or effect, and for  
CC improving gene expression. The present lentiviral vector system affords  
CC relatively high vector particle production, has improved transduction  
CC capabilities, and has even lower potential for replication competent  
CC retrovirus (RCR) events. This sequence represents woodchuck hepatitis  
CC virus posttranscriptional regulatory element (WPRE) used in the creation  
CC of the trans-lenti viral vector system.

XX SQ Sequence 592 BP; 77 A; 188 C; 148 G; 179 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 9; Length 592;  
Best Local Similarity 100.0%; Pred. No. 4.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTGGTGTG 60  
Db 328 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTGGTGTG 387

Qy 61 TCGGGGAAGCTGACGTCTTCCATGGCTGCTGCCCTGTGT 101

Db 388 TCGGGGAAGCTGACGTCTTCCATGGCTGCTGCCCTGTGT 428

## RESULT 2

AAD55110

ID AAD55110 standard; DNA; 604 BP.

XX AC AAD55110;

XX XX

XX 27-OCT-2003 (revised)

DT 07-AUG-2003 (first entry)

XX XX

XX Woodchuck hepatitis virus regulator element (WRE) DNA.

XX DE

XX Transgenic; biotechnology; Woodchuck hepatitis virus regulator element;

XX KW WRE; ds.

XX XX

XX Woodchuck hepatitis B virus.

XX OS

XX WO2003022040-A2.

XX PN

XX 20-MAR-2003.

XX PD

XX 12-SEP-2002; 2002WO-US029130.  
XX 13-SEP-2001; 2001US-0322031P.  
XX 09-JAN-2002; 2002US-0347782P.  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
XX PI Baltimore D, Hong EJ, Lois-Caballe C, Pease S;  
XX WPI; 2003-300976/29.

XX Producing a transgenic animal for commercial use, comprises transfecting  
XX a packaging cell line with retroviral construct, recovering recombinant  
XX retrovirus from the cell line and infecting embryonic cell with the  
XX recombinant virus.

XX Example 1; Fig 18B; 76pp; English.

XX The invention relates to a method for producing a transgenic animal for  
XX commercial use, which comprises transfecting a packaging cell line with  
XX retroviral construct, recovering recombinant retrovirus from the cell  
XX line and infecting embryonic cell with the recombinant virus. The method  
XX is useful in producing transgenic animals using retroviral constructs  
XX engineered to carry a transgene of interest. The transgenic animals may  
XX find use in commercial applications like biotechnology and agriculture.  
XX The present sequence is Woodchuck hepatitis virus regulator element (WRE)  
XX DNA used to illustrate the method of the invention. (Updated on 27-OCT-  
XX 2003 to standardise OS field)

XX SQ Sequence 604 BP; 81 A; 190 C; 150 G; 183 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 604;

Best Local Similarity 100.0%; Pred. No. 4.8e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTGGTGTG 60

Db 334 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTGGTGTG 393

Qy 61 TCGGGGAAGCTGACGTCTTCCATGGCTGCTGCCCTGTGT 101

Db 394 TCGGGGAAGCTGACGTCTTCCATGGCTGCTGCCCTGTGT 434

## RESULT 3

ACC45094

ID ACC45094 standard; DNA; 604 BP.

XX AC ACC45094;

XX XX

XX 27-OCT-2003 (revised)

DT 10-JUN-2003 (first entry)

XX XX

XX Woodchuck hepatitis virus regulator element (WRE) SEQ ID NO:4.

XX KW

XX Transgenic animal; transgenic bird; transgenic fish; transgene;

XX retroviral construct; lentiviral; long terminal repeat; LTR; WRE;

XX biotechnology; agriculture; woodchuck hepatitis virus; regulator element;

XX KW gene; ds.

XX OS

XX Woodchuck hepatitis B virus.

XX XX

XX WO2003022228-A2.

XX PN

XX 20-MAR-2003.

XX XX

XX 12-SEP-2002; 2002WO-US029157.

XX PF

XX 13-SEP-2001; 2001US-0322031P.

XX PR

XX 09-JAN-2002; 2002US-0347782P.

XX XX

XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.

XX Baltimore D, Hong EJ, Lois-Caballe C, Pease S;  
XX WPI; 2003-301005/29.  
XX Producing a transgenic bird or fish for commercial use, comprises  
XX transfecting a packaging cell line with retroviral construct, recovering  
XX recombinant retrovirus from the cell and infecting bird or fish egg with  
XX the recombinant virus.  
XX  
XX Example 1; Fig 18; 68pp; English.  
XX  
XX The present invention describes a method for producing a transgenic bird  
XX or fish. The method comprises transfecting a packaging cell line with a  
XX retroviral construct, recovering recombinant retroviral particles from  
XX the packaging cell line, and infecting a bird or a fish egg with the  
XX recombinant retroviral particles. The retroviral construct comprises the  
XX R and U5 sequences from a 5' lentiviral long terminal repeat (LTR) and a  
XX self-inactivating 3' lentiviral LTR. Also described is a transgenic bird  
XX or fish made by the above method and whose genome comprises a proviral  
XX DNA that has a self-inactivating 3' lentiviral LTR. The method is useful  
XX in producing transgenic animals, particularly transgenic birds and fish,  
XX using retroviral constructs engineered to carry a transgene of interest.  
XX The method is used to introduce the gene of choice into animals in order  
XX to confer upon them desired attributes. The transgenic animals may find  
XX use in commercial applications like biotechnology and agriculture. The  
XX present sequence represents a woodchuck hepatitis virus regulator element  
XX (WRE) nucleotide sequence, which is used in an example from the present  
XX invention. (Updated on 27-OCT-2003 to standardise OS field)  
XX  
XX Sequence 604 BP; 81 A; 190 C; 150 G; 183 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 101; DB 8; Length 604;  
XX Best Local Similarity 100.0%; Pred. No. 4.8e-22;  
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX Qy 1 TGCCTTCCCGCTCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60  
XX Db 334 TGCCTTCCCGCTCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 393  
XX  
XX Qy 61 TCGGGGAAGCTGACGTCTTCCATGGCTGCTCGCCTGTGT 101  
XX Db 394 TCGGGGAAGCTGACGTCTTCCATGGCTGCTCGCCTGTGT 434  
XX  
XX RESULT 4  
XX ADD67513  
XX ID ADD67513 standard; DNA; 609 BP.  
XX AC ADD67513;  
XX  
XX DT 29-JAN-2004 (first entry)  
XX  
XX DE WRE element #SEQ ID 1.  
XX  
XX Neuroprotective; antiparkinsonian; nootropic; anticonvulsant;  
XX transgene delivery; WRE element; APP5'UTR; tau3'UTR; TH3'UTR; vector;  
XX neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
XX amyotrophic lateral sclerosis; Huntington's disease;  
XX retinal degenerative disease; posttranscription; ds.  
XX  
XX OS Synthetic.  
XX  
XX PN EPI361277-A1.  
XX  
XX PD 12-NOV-2003.  
XX  
XX PF 30-APR-2002; 2002EP-00291091.  
XX  
XX PR 30-APR-2002; 2002EP-00291091.  
XX  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX (BIOV-) BIOVECTYS.  
XX  
XX PI Kim J;  
XX  
XX Mallet J, Brun S, Dufour N, Faucon-Biguet N;  
XX WPI; 2003-879907/82.  
XX  
XX New vector, for transgene delivery into mammalian cells, comprising a  
XX chimeric genetic construct with a transgene linked to a WPRE element, or  
XX APP5'UTR, tau3'UTR or TH3'UTR region, useful for treating  
XX neurodegenerative disease.  
XX  
XX Claim 9; SEQ ID NO 1; 30pp; English.  
XX  
XX The invention relates to a vector for transgene delivery into mammalian  
XX cells. The vector comprises a chimeric genetic construct with a transgene  
XX operably linked to at least two distinct posttranscriptional regulatory  
XX elements, e.g. WPRE element, APP5'UTR, tau3'UTR or TH3'UTR region. The  
XX WPRE element, APP5'UTR, tau3'UTR and TH3'UTR region or their functional  
XX fragment comprises the nucleotide sequence of 609, 95, 237 and 91 bp,  
XX respectively ADD67513-ADD67516. The vector comprises a promoter  
XX controlling transcription of the transgene in the mammalian cells, a  
XX marker gene and a polyadenylation signal operably linked to the  
XX transgene. The vector is a plasmid or a recombinant virus. The vector or  
XX recombinant cell is used for the manufacture of a medicament for treating  
XX a disease. The vector, recombinant cell or composition is useful for  
XX treating a human disease, e.g. neurodegenerative diseases selected from  
XX Parkinson's disease, Alzheimer's disease, amyotrophic lateral sclerosis,  
XX Huntington's disease or retinal degenerative diseases. They can also be  
XX used in experiments research or prophylactic areas. The current sequence  
XX represents the WPRE element nucleotide sequence.  
XX  
XX Sequence 609 BP; 83 A; 191 C; 151 G; 184 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 101; DB 10; Length 609;  
XX Best Local Similarity 100.0%; Pred. No. 4.8e-22;  
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX Qy 1 TGCCTTCCCGCTCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 60  
XX Db 338 TGCCTTCCCGCTCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTG 397  
XX  
XX Qy 61 TCGGGGAAGCTGACGTCTTCCATGGCTGCTCGCCTGTGT 101  
XX Db 398 TCGGGGAAGCTGACGTCTTCCATGGCTGCTCGCCTGTGT 438  
XX  
XX RESULT 5  
XX ADO25310  
XX ID ADO25310 standard; DNA; 632 BP.  
XX AC ADO25310;  
XX  
XX DT 12-AUG-2004 (first entry)  
XX  
XX DE Woodchuck hepatitis virus WPRE sequence as biological cloning marker.  
XX  
XX ss; porcine; uroplakin II gene; promoter; expression vector;  
XX surrogate mother animal; transgenic animal; urine; bladder.  
XX  
XX OS Woodchuck hepatitis B virus.  
XX  
XX PN WC2004042062-A1.  
XX  
XX PD 21-MAY-2004.  
XX  
XX PF 04-NOV-2003; 2003WO-KR002339.  
XX  
XX PR 04-NOV-2002; 2002KR-00067856.  
XX  
XX PR 03-NOV-2003; 2003KR-00077256.  
XX  
XX PA (CHOA-) CHO-A PHARM CO LTD.  
XX (KIMJ/) KIM J.  
XX  
XX PI Kim J;  
XX

XX WPI; 2004-411520/38.  
XX  
XX Novel porcine uroplakin II gene promoter, useful for promoting the  
XX bladder-specific expression of a specific target protein.  
XX  
XX Claim 7; SEQ ID NO 7; 76pp; English.  
XX  
XX The invention relates to a novel porcine uroplakin II gene promoter (I).  
XX An expression vector containing the promoter is useful in a method of  
XX producing useful proteins which involves implanting the vector into a  
XX surrogate mother animal, obtaining transgenic animals from the surrogate  
XX mother animal, and isolating and purifying useful proteins from the urine  
XX of the transgenic animals. The promoter, the expression vector, and the  
XX transgenic animal can be used in the production field of useful proteins  
XX that are medicinally valuable. The promoter promotes the bladder-specific  
XX expression of a target protein at high efficiency. An animal which was  
XX transformed using the promoter, so as to express the target protein,  
XX secret with a target protein in its urine at high concentration, and the  
XX protein thus produced shows a superior physiological activity to that of  
XX the same kind of the existing protein. This sequence represents the  
XX woodchuck hepatitis virus posttranslational regulatory element (WPRE)  
XX sequence which is used as a marker gene for cloning of genes of interest  
XX and linking to the porcine uroplakin II gene promoter.  
XX  
XX Sequence 632 BP; 80 A; 201 C; 158 G; 193 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 101; DB 12; Length 632;  
Best Local Similarity 100.0%; Pred. No. 4.9e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGT 60  
DB 347 TGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGT 406  
QY 61 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTCGCCTGTGT 101  
DB 407 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTCGCCTGTGT 447  
RESULT 6  
AAD55114  
ID AAD55114 standard; DNA; 2853 BP.  
AC AAD55114;  
XX  
XX 07-AUG-2003 (first entry)  
XX  
XX HflapubigWRE chimeric construct DNA.  
XX  
XX Transgenic; biotechnology; agriculture; green fluorescent protein; GFP;  
XX Woodchuck hepatitis virus regulator element; WRE; human; ubiquitin;  
XX Human immunodeficiency virus type 1; HIV-1; chimeric; ds.  
XX  
XX Human immunodeficiency virus 1.  
XX Unidentified.  
XX OS Homo sapiens.  
XX OS Woodchuck hepatitis B virus.  
XX OS Chimeric.  
XX  
XX WO2003022040-A2.  
XX  
XX 20-MAR-2003.  
XX  
XX 12-SEP-2002; 2002WO-US029130.  
XX  
XX 13-SEP-2001; 2001US-0322031P.  
XX 09-JAN-2002; 2002US-0347782P.  
XX  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
XX  
XX Baltimore D, Hong EJ, Lois-Caballe C, Pease S;  
XX  
XX WPI; 2003-301005/29.  
XX  
XX Producing a transgenic bird or fish for commercial use, comprises  
XX transfecting a packaging cell line with retroviral construct, recovering

DR WPI; 2003-300976/29.  
XX  
XX Producing a transgenic animal for commercial use, comprises transfecting  
XX a packaging cell line with retroviral construct, recovering recombinant  
XX retrovirus from the cell line and infecting embryonic cell with the  
XX recombinant virus.  
XX  
XX Example 1; Fig 21; 76pp; English.  
XX  
XX The invention relates to a method for producing a transgenic animal for  
XX commercial use, which comprises transfecting a packaging cell line with  
XX retroviral construct, recovering recombinant retrovirus from the cell  
XX line and infecting embryonic cell with the recombinant virus. The method  
XX is useful in producing transgenic animals using retroviral constructs  
XX engineered to carry a transgene of interest. The transgenic animals may  
XX find use in commercial applications like biotechnology and agriculture.  
XX The present sequence is HflapubigWRE chimeric construct DNA comprising  
XX Human immunodeficiency virus type 1 (HIV-1) NL4.3 flap sequence, green  
XX fluorescent protein (GFP) variant encoding sequence, human ubiquitin  
XX promoter sequence and Woodchuck hepatitis virus regulator element (WRE).  
XX This sequence is used to illustrate the method of the invention  
SQ  
Sequence 2853 BP; 564 A; 774 C; 858 G; 657 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 8; Length 2853;  
Best Local Similarity 100.0%; Pred. No. 6.5e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGT 60  
DB 2559 TGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGT 2618  
QY 61 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTCGCCTGTGT 101  
DB 2619 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTCGCCTGTGT 2659  
RESULT 7  
ACC45098  
ID ACC45098 standard; DNA; 2853 BP.  
XX  
XX ACC45098;  
XX  
XX 10-JUN-2003 (first entry)  
XX  
XX  
XX HIV-1 flap + ubiquitin + GFP + WRE construct DNA sequence SEQ ID NO:8.  
XX  
XX Transgenic animal; transgenic bird; transgenic fish; transgene;  
XX retroviral construct; lentiviral; long terminal repeat; LTR;  
XX biotechnology; agriculture; gene; ds.  
XX  
XX Human immunodeficiency virus 1.  
XX OS Homo sapiens.  
XX OS Woodchuck hepatitis B virus.  
XX OS Synthetic.  
XX  
XX WO2003022228-A2.  
XX  
XX 20-MAR-2003.  
XX  
XX 12-SEP-2002; 2002WO-US029157.  
XX  
XX 13-SEP-2001; 2001US-0322031P.  
XX 09-JAN-2002; 2002US-0347782P.  
XX  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
XX  
XX Baltimore D, Hong EJ, Lois-Caballe C, Pease S;  
XX  
XX WPI; 2003-301005/29.  
XX  
XX Producing a transgenic bird or fish for commercial use, comprises  
XX transfecting a packaging cell line with retroviral construct, recovering

PT recombinant retrovirus from the cell and infecting bird or fish egg with  
PT the recombinant virus.  
XX  
XX Example 1; Fig 21; 68pp; English.  
XX  
XX The present invention describes a method for producing a transgenic bird  
XX or fish. The method comprises transfecting a packaging cell line with a  
XX retroviral construct, recovering recombinant retroviral particles from  
XX the packaging cell line, and infecting a bird or a fish egg with the  
XX recombinant retroviral particles. The retroviral construct comprises the  
XX R and U5 sequences from a 5' lentiviral long terminal repeat (LTR) and a  
XX self-inactivating 3' lentiviral LTR. Also described is a transgenic bird  
XX or fish made by the above method and whose genome comprises a proviral  
XX DNA that has a self-inactivating 3' lentiviral LTR. The method is useful  
XX in producing transgenic animals, particularly transgenic birds and fish,  
XX using retroviral constructs engineered to carry a transgene of interest.  
XX The method is used to introduce the gene of choice into animals in order  
XX to confer upon them desired attributes. The transgenic animals may find  
XX use in commercial applications like biotechnology and agriculture. The  
XX present sequence represents a construct nucleotide sequence comprising an  
XX HIV-1 sequence, a green fluorescent protein (GFP) variant sequence, a  
XX human ubiquitin promoter sequence and a woodchuck hepatitis regulator  
XX element sequence, which is used in an example from the present invention  
XX  
SQ Sequence 2853 BP; 564 A; 774 C; 858 G; 657 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 101; DB 8; Length 2853;  
Best Local Similarity 100.0%; Pred. No. 6.5e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAAATTCGGTGTGTTG 60  
Db 2559 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAAATTCGGTGTGTTG 2618  
  
QY 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 101  
Db 2619 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 2659  
  
RESULT 8  
ABV77010  
ID ABV77010 standard; DNA; 7515 BP.  
XX  
XX ABV77010;  
AC  
XX  
XX 03-MAR-2003 (first entry)  
DT  
XX  
XX Nucleotide sequence of plasmid pSmart2 5'cppt.  
DE  
XX  
XX Viral vector; adipose tissue; adipose tissue metabolism; obesity;  
KW diabetes; blood disorder; vascular disease; ss.  
XX  
XX Synthetic.  
OS  
XX  
XX WO200286132-A2.  
FN  
XX  
XX 31-OCT-2002.  
PD  
XX  
XX 19-APR-2002; 2002WO-GB001830.  
PP  
XX  
XX 20-APR-2001; 2001GB-00009781.  
PR  
XX  
XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
PA  
XX  
XX Kingsman SA, Mitrophanous K, Ellard FM;  
PI  
XX  
XX WPI; 2003-093139/08.  
DR  
XX  
XX Use of viral vector system for transducing a target adipose tissue site,  
PT and for treating and/or preventing vascular diseases or diseases  
PT associated with death or impaired function of adipose tissue cells, such  
PT as obesity and diabetes.  
XX

PS Example; Page 65-67; 84pp; English.  
XX  
XX The specification describes the use of a viral vector system for  
XX transducing a target adipose tissue site. The viral vector system is  
XX useful for transducing a target adipose tissue site, in the manufacture  
XX of a pharmaceutical composition for treating and/or preventing a disease  
XX associated with a derangement in the metabolism of adipose tissue, such  
XX as obesity and diabetes. The viral vector system is also useful for  
XX treating and preventing a disease associated with death or impaired  
XX function of adipose tissue cells, a disease associated with hereditary  
XX blood disorders, and vascular diseases. The present sequence represents  
XX plasmid pSmart 5'cppt, which is used to construct vectors for use in the  
XX invention  
SQ Sequence 7515 BP; 1978 A; 1733 C; 1829 G; 1975 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 101; DB 8; Length 7515;  
Best Local Similarity 100.0%; Pred. No. 7.8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAAATTCGGTGTGTTG 60  
Db 4424 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAAATTCGGTGTGTTG 4483  
  
QY 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 101  
Db 4484 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 4524  
  
RESULT 9  
ADM47497  
ID ADM47497 standard; DNA; 7648 BP.  
XX  
XX ADM47497;  
AC  
XX  
XX 03-JUN-2004 (first entry)  
DT  
XX  
XX Lysosomal enzyme related DNA #1.  
DE  
XX  
XX Lysosomal enzyme; lysosomal storage disease; Gaucher type I disease;  
KW Hurler disease; Sanfilippo disease; gene therapy; ds..  
XX  
XX Unidentified.  
OS  
XX  
XX US2004023218-A1.  
FN  
XX  
XX 05-FEB-2004.  
PD  
XX  
XX 21-JUN-2002; 2002US-00176066.  
PP  
XX  
XX 21-JUN-2002; 2002US-00176066.  
PR  
XX  
XX (DESM/) DESMARIS N.  
PA  
XX  
XX (HEAR/) HEARD J M.  
PI  
XX  
XX Desmaris N, Heard JM;  
PI  
XX  
XX WPI; 2004-142648/14.  
DR  
XX  
XX New purified nucleic acid molecules capable of expressing a lysosomal  
PT enzyme, useful for preventing or treating lysosomal storage diseases  
PT (e.g. Gaucher type I disease or Sanfilippo disease) in humans.  
XX  
XX Disclosure; Page 9-11; 20pp; English.  
XX  
XX The invention relates to a nucleic acid molecule capable of expressing a  
XX lysosomal enzyme. The nucleic acid molecule comprises at least a sequence  
XX coding for the lysosomal enzyme and a promoter highly active in the brain  
XX inserted upstream from the sequence. Compositions and methods of the  
XX invention are useful for preventing or treating lysosomal storage  
XX diseases e.g. Gaucher type I disease, Hurler disease (MPSI) or Sanfilippo  
XX disease (MPSIII) in humans. The invention is also useful in gene therapy.  
XX The present sequence is a DNA related to the invention.  
XX

XX SQ Sequence 7648 BP; 1546 A; 2244 C; 2052 G; 1806 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 7648;  
Best Local Similarity 100.0%; Pred. No. 7.9e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60  
|||||  
Db 3303 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 3362  
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101  
|||||  
Db 3363 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 3403  
|||||

RESULT 10  
ADMA47498  
ID ADMA47498 standard; DNA; 8092 BP.  
XX  
XX ADMA47498;  
AC  
XX  
XX 03-JUN-2004 (first entry)  
DT  
XX  
XX Lysosomal enzyme related DNA #2.  
DE  
XX  
XX Lysosomal enzyme; lysosomal storage disease; Gaucher type I disease;  
KW Hurler disease; Sanfilippo disease; gene therapy; ds.  
XX  
XX Unidentified.  
OS  
XX US2004023218-A1.  
PN  
XX  
XX 05-FEB-2004.  
PD  
XX  
XX 21-JUN-2002; 2002US-00176066.  
PF  
XX  
XX 21-JUN-2002; 2002US-00176066.  
PR  
XX  
XX (DESM/) DESMARIS N.  
PA (HEAR/) HEARD J M.  
XX  
XX Desmaris N, Heard JM;  
PI  
XX WPI; 2004-142648/14.  
DR  
XX  
XX New purified nucleic acid molecules capable of expressing a lysosomal  
PT enzyme, useful for preventing or treating lysosomal storage diseases  
PT (e.g. Gaucher type I disease or Sanfilippo disease) in humans.  
XX  
XX Disclosure; Page 11-15; 20pp; English.

XX SQ The invention relates to a nucleic acid molecule capable of expressing a  
CC lysosomal enzyme. The nucleic acid molecule comprises at least a sequence  
CC coding for the lysosomal enzyme and a promoter highly active in the brain  
CC inserted upstream from the sequence. Compositions and methods of the  
CC invention are useful for preventing or treating lysosomal storage  
CC diseases e.g. Gaucher type I disease, Hurler disease (MPSI) or Sanfilippo  
CC disease (MPSIII) in humans. The invention is also useful in gene therapy.  
XX The present sequence is a DNA related to the invention.

XX SQ Sequence 8092 BP; 1650 A; 2268 C; 2226 G; 1948 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 8092;  
Best Local Similarity 100.0%; Pred. No. 7.9e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60  
|||||  
Db 3747 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 3806  
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101  
|||||

Db 3807 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 3847

RESULT 11  
AAA59091  
ID AAA59091 standard; DNA; 8484 BP.  
XX  
XX AAA59091;  
AC  
XX  
XX 07-NOV-2000 (first entry)  
DT  
XX  
XX Nucleotide sequence of plasmid pDV90.  
DE  
XX  
XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;  
KW ss.  
XX  
XX Synthetic.  
OS  
XX WO200042208-A1.  
PN  
XX  
XX 20-JUL-2000.  
PD  
XX  
XX 14-JAN-2000; 2000WO-EP000265.  
PF  
XX  
XX 14-JAN-1999; 99US-0115920P.  
PR  
XX  
XX (NOVS ) NOVARTIS AG.  
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
PA (SCRI ) SCRIPPS RES INST.  
XX  
XX Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;  
PI Skripchenko Y;  
PI  
XX WPI; 2000-476068/41.  
DR  
XX  
XX New nucleic acid comprising an adenovirus tripartite leader nucleotide  
PT for producing high-capacity and targeted vectors for adenovirus-based  
PT gene therapy.

XX SQ Claim 10; Page 200-204; 212pp; English.

XX SQ The specification describes a nucleic acid molecule comprising an  
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence  
CC comprising two different TPL exons or three same or different TPL exons.  
CC The nucleic acid is used to produce an adenovirus vector particle,  
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral  
CC vectors, target an adenovirus vector to a cell, produce a modified  
CC adenovirus, deliver a heterologous gene to an animal and produce a  
CC gutless adenoviral vector particle. The present sequence represents  
CC plasmid pDV90, which contains a TPL

XX SQ Sequence 8484 BP; 1996 A; 2238 C; 2125 G; 2125 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 3; Length 8484;  
Best Local Similarity 100.0%; Pred. No. 8e-22;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 60  
|||||  
Db 4192 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGTG 4251  
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101  
|||||

Db 4252 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 4292  
|||||

RESULT 12  
ADF48775  
ID ADF48775 standard; DNA; 8484 BP.  
XX  
XX ADF48775;  
AC  
XX  
XX 12-FEB-2004 (first entry)  
DT





Db 6513 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 6572

Qy 61 TCGGGGAAGCTGACGTCTTCCATGCTGCTGCTGCCCTGTGT 101  
 Db 6573 TCGGGGAAGCTGACGTCTTCCATGCTGCTGCTGCCCTGTGT 6613

RESULT 14  
 ADM82791/c  
 ID ADM82791 standard; cDNA; 9731 BP.  
 XX AC ADM82791;  
 XX AC  
 XX DT 03-JUN-2004 (first entry)  
 XX DE  
 XX DE DNA repair pathway related retroviral vector cDNA with CMV promoter.  
 KW inducer; inhibitor; DNA repair pathway; anti-HIV; cytostatic; virucide;  
 KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;  
 KW cancer; human adult T-cell leukaemia; lymphoma;  
 KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;  
 KW gene therapy; human; cyclic; circular; CMV promoter; ss.  
 XX OS Unidentified.  
 XX OS  
 XX PN WO2003089573-A2.  
 XX PN  
 XX PD 30-OCT-2003.  
 XX PD  
 XX PF 04-APR-2003; 2003WO-US010302.  
 XX PF  
 XX PR 05-APR-2002; 2002US-0370376P.  
 XX PR  
 XX PA (FISH/) FISHEL R A.  
 XX PA (YODE/) YODER K E.  
 XX PI Fishel RA, Yoder KE;  
 XX PI  
 XX DR WPI; 2003-854096/79.  
 XX DR  
 XX PT Screening for compounds that modulate a DNA repair pathway and/or  
 PT retroviral integration, useful for treating retroviral infection,  
 PT comprises determining the amount of a retroviral cDNA circularization in  
 PT the presence of the test compound.  
 XX PS  
 XX PS Claim 73; SEQ ID NO 5; 89pp; English.  
 XX CC The invention relates to a novel method for screening for inducers or  
 CC inhibitors of a DNA repair pathway by contacting at least one component  
 CC of a DNA repair pathway with a non-circularized retroviral cDNA in the  
 CC presence and absence of a test compound, and determining whether  
 CC circularization of the cDNA is increased or decreased in the presence of  
 CC the test compound. The DNA repair pathway components have the following  
 CC activities: anti-HIV, cytostatic, virucide, antidiabetic, and  
 CC neuroprotective. The method is useful for identifying compounds that  
 CC modulate a DNA repair pathway and/or retroviral activity. The compound is  
 CC used in manufacturing a pharmaceutical composition for the treatment of a  
 CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-  
 CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes  
 CC or multiple sclerosis) or for increasing the efficiency of gene delivery  
 CC in a gene therapy. This polynucleotide represents a retroviral cDNA  
 CC sequence of the invention.  
 XX SQ  
 SQ Sequence 9731 BP; 2444 A; 2412 C; 2548 G; 2327 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 11; Length 9731;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-22;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60  
 Db 7288 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 7229

Qy 61 TCGGGGAAGCTGACGTCTTCCATGCTGCTGCTGCCCTGTGT 101  
 Db 7228 TCGGGGAAGCTGACGTCTTCCATGCTGCTGCTGCCCTGTGT 7188

RESULT 15  
 ADM82792/c  
 ID ADM82792 standard; cDNA; 9782 BP.  
 XX AC ADM82792;  
 XX AC  
 XX DT 03-JUN-2004 (first entry)  
 XX DE  
 XX DE DNA repair pathway related retroviral vector cDNA with MSH2 promoter.  
 KW inducer; inhibitor; DNA repair pathway; anti-HIV; cytostatic; virucide;  
 KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;  
 KW cancer; human adult T-cell leukaemia; lymphoma;  
 KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;  
 KW gene therapy; human; cyclic; circular; MSH2 promoter; ss.  
 XX OS Unidentified.  
 XX OS  
 XX PN WO2003089573-A2.  
 XX PN  
 XX PD 30-OCT-2003.  
 XX PD  
 XX PF 04-APR-2003; 2003WO-US010302.  
 XX PF  
 XX PR 05-APR-2002; 2002US-0370376P.  
 XX PR  
 XX PA (FISH/) FISHEL R A.  
 XX PA (YODE/) YODER K E.  
 XX PI Fishel RA, Yoder KE;  
 XX PI  
 XX DR WPI; 2003-854096/79.  
 XX DR  
 XX PT Screening for compounds that modulate a DNA repair pathway and/or  
 PT retroviral integration, useful for treating retroviral infection,  
 PT comprises determining the amount of a retroviral cDNA circularization in  
 PT the presence of the test compound.  
 XX PS  
 XX PS Claim 73; SEQ ID NO 6; 89pp; English.  
 XX CC The invention relates to a novel method for screening for inducers or  
 CC inhibitors of a DNA repair pathway by contacting at least one component  
 CC of a DNA repair pathway with a non-circularized retroviral cDNA in the  
 CC presence and absence of a test compound, and determining whether  
 CC circularization of the cDNA is increased or decreased in the presence of  
 CC the test compound. The DNA repair pathway components have the following  
 CC activities: anti-HIV, cytostatic, virucide, antidiabetic, and  
 CC neuroprotective. The method is useful for identifying compounds that  
 CC modulate a DNA repair pathway and/or retroviral activity. The compound is  
 CC used in manufacturing a pharmaceutical composition for the treatment of a  
 CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-  
 CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes  
 CC or multiple sclerosis) or for increasing the efficiency of gene delivery  
 CC in a gene therapy. This polynucleotide represents a retroviral cDNA  
 CC sequence of the invention.  
 XX SQ  
 SQ Sequence 9782 BP; 2628 A; 2305 C; 2351 G; 2498 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 11; Length 9782;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-22;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60  
 Db 1340 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 1281

Qy 61 TCGGGGAAGCTGACGTCTTCCATGCTGCTGCTGCCCTGTGT 101

Db 1280 TCGGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 1240

Search completed: July 14, 2005, 07:01:51  
Job time : 145.448 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	101	100.0	142	6	AR356490	AR356490 Sequence	
C 2	101	100.0	142	6	AR538046	AR538046 Sequence	
C 3	101	100.0	228	6	E00019	E00019 DNA coding	
C 4	101	100.0	240	1	PM0000	PM0000 DNA coding	
C 5	101	100.0	251	6	E00018	E00018 DNA coding	
C 6	101	100.0	251	6	I01644	I01644 Sequence 1	
C 7	101	100.0	344	11	HUMUT5345	Li18624 Human chromosome 5	
C 8	101	100.0	400	6	BD195256	BD195256 Nucleotide	
C 9	101	100.0	456	6	E00892	E00892 Synthetic D	
C 10	101	100.0	456	6	E01156	E01156 DNA fragment	
C 11	101	100.0	456	6	E01274	E01274 DNA encoding	
C 12	101	100.0	456	6	E01302	E01302 DNA encoding	
C 13	101	100.0	466	6	AX260098	AX260098 Sequence	
C 14	101	100.0	573	6	AX260150	AX260150 Sequence	
C 15	101	100.0	693	6	A43586	A43586 Sequence 11	
C 16	101	100.0	693	6	AR116755	AR116755 Sequence	
C 17	101	100.0	998	1	AY559171	AY559171 Pseudomon	
C 18	101	100.0	1011	1	SMTEA068	X97254 S.marcescens	
C 19	101	100.0	1012	2	CEC11F10	Z92776 Caenorhabdit	

RESULT 2	AR538046/c	AR538046	142 bp	DNA	
LOCUS		Sequence	2608	from patent	US 6'37248.
DEFINITION		AR538046			
ACCESSION		AR538046			
VERSION		AR538046.1	GI:53929263		
KEYWORDS					
SOURCE					Unknown.

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES
    source
        Location/Qualifiers
            1..142
                /organism="unknown"
                /mol_type="genomic DNA"
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 142;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 48
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 228)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FT CDS 210..>228
    /product='E.coli penicillinase'.
FEATURES
    source
        Location/Qualifiers
            1..228
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /db_xref="taxon:562"
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 228;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 116
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101

```

```

Db 115 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75
RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences, plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES
    source
        Location/Qualifiers
            1..240
                /organism="Plasmid pMM110"
                /mol_type="genomic DNA"
                /db_xref="taxon:2599"
                /plasmid="Plasmid pMM110"
ORIGIN Unreported.
    Query Match 100.0%; Score 101; DB 1; Length 240;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 92
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 251)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: Clone=pKT241;

```

FEATURES  
source  
1..251  
Location/Qualifiers  
/organism="Escherichia coli"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:562"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 251;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60  
|||||  
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116  
|||||

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
|||||  
DB 115 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 75  
|||||

RESULT 6  
I01644/c  
LOCUS 251 bp ss-DNA linear PAT 18-MAY-1993  
DEFINITION Sequence 1 from Patent US 4338397.  
ACCESSION I01644  
VERSION I01644.1 GI:267685  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 251)  
AUTHORS Gilbert, W. and Talmadge, K.  
TITLE Mature protein synthesis  
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;  
President and Fellows of Harvard College; Cambridge, MA

FEATURES  
source  
1..251  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 251;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60  
|||||  
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116  
|||||

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
|||||  
DB 115 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 75  
|||||

RESULT 7  
HUMUT5345  
LOCUS 344 bp DNA linear STS 26-JUL-1993  
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.  
ACCESSION L18624  
VERSION L18624.1 GI:308338  
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;  
microsatellite repeat; repeat polymorphism; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 344)  
AUTHORS Gerken, S.C., Matsunami, N., Lawrence, E., Carlson, M., Moore, M.,

TITLE  
JOURNAL  
COMMENT  
Ballard, L., Melis, R., Robertson, M., Bradley, P., Elsner, T.,  
Tingey, A., Rodriguez, P., Albertsen, H., Lalouel, J.-M. and White, R.  
Genetic and physical mapping of simple sequence repeat containing  
sequence tagged sites from the human genome  
Unpublished (1993)  
Original source text: Homo sapiens DNA.  
Submitted by: Utah Center for Human Genome Research University of  
Utah, Dept. of Human Genetics  
2160 Eccles Institute of Human Genetics  
Salt Lake City, UT 84112  
e-mail: sts@corona.med.utah.edu  
Primer A: GAGCAAAACAGGAGGCAAAATGC  
Primer B: TTCGGGAATGTGCGGAAC  
32P-label: B Primer  
PCR Profile:  
Initial Denaturation: 94C 300sec  
PCR Cycles: 30  
Denaturation: 94C 10sec  
Annealing: 60C 10sec  
Extension: 72C 20sec  
Mg++: 2mM  
Gel: Acrylamide 7%, Formamide 32%, Urea 34%  
Alleles: 2  
Location/Qualifiers  
1..344  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="8"  
36..224  
/standard\_name="STS UT5345"  
36..60  
complement(202..224)

STS  
primer\_bind  
primer\_bind  
ORIGIN  
Query Match 100.0%; Score 101; DB 11; Length 344;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60  
|||||  
DB 141 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 200  
|||||

QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
|||||  
DB 201 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 241  
|||||

RESULT 8  
BD195256/c  
LOCUS 400 bp DNA linear PAT 17-JUL-2003  
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.  
ACCESSION BD195256  
VERSION BD195256.1 GI:33005021  
KEYWORDS JP 2002513277-A/43.  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.  
REFERENCE 1 (bases 1 to 400)  
AUTHORS Dillon, P.J., Choi, G.H. and Welch, R.A.  
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands  
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;  
HUMAN GENOME SCIENCES INC, WISCONSIN ALUMNI RESEARCH FOUNDATION  
COMMENT OS Unidentified  
PN JP 2002513277-A/43  
PD 08-MAY-2002  
PF 21-NOV-1997 JP 1998523916  
PR 22-NOV-1996 US 60/031626, 14-OCT-1997 US 60/061953 PI  
PATRICK J DILLON, GIL H CHOI, RODNEY A WELCH  
PC C12N15/11, C12N15/63, C07K16/12, G01N33/569, G06F17/30, G11B7/00 CC  
Strandedness: Double;  
CC Topology: Linear;  
CC Nucleotide sequence of Escherichia coli pathogenicity islands

FEATURES  
source  
FH Key Location/Qualifiers  
FT source 1..400  
FT /organism='Unidentified'.  
1..400  
/organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 400;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60  
Db 165 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 106  
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
Db 105 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 65

RESULT 9  
E00892/c  
LOCUS  
DEFINITION Synthetic DNA encoding fused polypeptide between E coli  
E00892  
ACCESSION E00892.1 GI:2169153  
VERSION JP 1986149089-A/1  
KEYWORDS JP 1986149089-A/1  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,  
Ojida,K. and Matsushiro,A.  
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED  
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM  
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;  
EARTH CHEM CORP LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
FN JP 1986149089-A/1  
PD 07-JUL-1986  
PF 21-DEC-1984 JP 1984271206  
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,  
PI TOCHIFUSA NORIYUKI,  
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC  
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC  
C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: strain=HB101;  
CC \*source: clone=pVG201;  
CC Feature is identified by experimental;  
FH Key Location/Qualifiers  
FH promoter 125..170  
FT of beta-lactamase  
FT RBS 200..203  
FT CDS 209..438  
FT sig\_peptide 209..277  
FT mat\_peptide 278..435  
FT /product='beta-urogastrone precursor' FT  
FT /product='signal peptide of beta-lactonase' FT  
FT /product='mature peptide'.  
FT Location/Qualifiers  
1..456  
/organism='synthetic construct'  
/mol\_type='genomic DNA'

FEATURES  
source  
FH Key Location/Qualifiers  
FH promoter 125..170  
FT of beta-lactamase  
FT RBS 200..203  
FT CDS 209..438  
FT sig\_peptide 209..277  
FT mat\_peptide 278..435  
FT /product='beta-urogastrone precursor' FT  
FT /product='signal peptide of beta-lactonase' FT  
FT /product='mature peptide'.  
FT Location/Qualifiers  
1..456  
/organism='synthetic construct'  
/mol\_type='genomic DNA'

ORIGIN  
/db\_xref='taxon:32630'

Query Match 100.0%; Score 101; DB 6; Length 456;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60  
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 114  
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
Db 113 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73

RESULT 10  
E01156/c  
LOCUS  
DEFINITION DNA fragment which secrets beta urogastrone.  
E01156  
ACCESSION E01156.1 GI:2169415  
VERSION JP 1987083890-A/1  
KEYWORDS JP 1987083890-A/1  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.  
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED  
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID  
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;  
EARTH CHEM CORP LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
FN JP 1987083890-A/1  
PD 17-APR-1987  
PF 09-OCT-1985 JP 1985225393  
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI  
KOIDE TAKAO,  
PI OKAI HIDEO  
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC  
C12R1:125);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: clone=pUG201;  
FH Key Location/Qualifiers  
FH promoter 125..170  
FT /note='beta lactamase promoter' FT RBS  
FT CDS 209..439  
FT /product='beta urogastrone'  
FT sig\_peptide 209..277  
FT mat\_peptide 278..436  
FT /product='beta urogastrone'.  
FT Location/Qualifiers  
1..456  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 456;  
Best Local Similarity 100.0%; Pred. No. 8.7e-20;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60  
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 114  
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101



```
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS E01274 DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsuura,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaiharu,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'
FEATURES
source Location/Qualifiers
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION E01302.1 GI:2169561
KEYWORDS JP 1987190083-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsuura,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaiharu,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL Patent: JP 1987190083-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'
FEATURES
source Location/Qualifiers
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source Location/Qualifiers
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
```

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 221
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||||
Db 220 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180
    |||||||

RESULT 14
AX260150/c
LOCUS      AX260150          573 bp      DNA      linear      PAT 26-OCT-2001
DEFINITION Sequence 112 from Patent WO0172774.
ACCESSION  AX260150
VERSION     AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM   Drosophila melanogaster (fruit fly)
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1
AUTHORS    Deak, P., Glover, D.M. and Midgley, C.
TITLE      Cell cycle progression proteins
JOURNAL    Patent: WO 0172774-A 112 04-OCT-2001;
            Cyclacel Limited (GB)
FEATURES   source
            1..573
            /organism="Drosophila melanogaster"
            /mol_type="unassigned DNA"
            /db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 296
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||||
Db 295 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 255
    |||||||

RESULT 15
A43586
LOCUS      A43586          693 bp      DNA      linear      PAT 06-MAR-1997
DEFINITION Sequence 11 from Patent WO9507357.
ACCESSION  A43586
VERSION     A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM   Cuphea lanceolata
            Cuphea lanceolata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; Myrtales; Lythraceae; Cuphea.
            1 (bases 1 to 693)
REFERENCE  1
AUTHORS    Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
            Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
            Schulte, W., Voetz, M., Walek, J. and Schell, J.
            PROMOTERS
TITLE      Patent: WO 9507357-A 11 16-MAR-1995;
            MAX PLANCK GESELLSCHAFT (DE)
JOURNAL
COMMENT    Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 651
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||||
Db 652 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 692
    |||||||

Search completed: July 14, 2005, 14:03:38
Job time : 757.618 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_8384\_8484

Perfect score: 101  
Sequence: 1 aggtttattgtctcatgacg.....gaaaagtgcacacgcagtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2	AAV76919 Staphyloc
C 2	101	100.0	228	1	AAN10032 Sequence
C 3	101	100.0	251	1	AAN10031 Sequence
C 4	101	100.0	400	2	AAV31229 E. coli J
C 5	101	100.0	456	1	AAN60624 Plasmid p
C 6	101	100.0	456	1	AAN71080 Sequence
C 7	101	100.0	456	1	AAN70833 Beta-urog
C 8	101	100.0	456	1	AAN81765 Sequence
C 9	101	100.0	466	6	ABA90413 Drosophil
C 10	101	100.0	487	2	AAx21173 Polynucle
C 11	101	100.0	535	2	AAx21149 Polynucle
C 12	101	100.0	573	6	ABA90456 Drosophil
C 13	101	100.0	605	12	ADH58311 Electroph
C 14	101	100.0	776	4	AAx30560 DNA encod
C 15	101	100.0	776	4	AAS27819 DNA encod
C 16	101	100.0	776	4	ABK42984 Genomic s
C 17	101	100.0	776	4	AAI07344 Human rep
C 18	101	100.0	776	4	AAI03229 Human rep
C 19	101	100.0	776	4	AAI06588 Human rep
C 20	101	100.0	776	4	AAI07340 Human rep

C 21	101	100.0	776	5	ABA14573 Human ner
C 22	101	100.0	776	5	AAS34681 Human DNA
C 23	101	100.0	776	8	ADA41574 Human sec
C 24	101	100.0	776	8	ACC50905 Human sec
C 25	101	100.0	776	8	ABZ71508 Secreted
C 26	101	100.0	776	9	ADB91869 Human sec
C 27	101	100.0	776	9	ADB61140 Connectiv
C 28	101	100.0	776	10	ADB94622 Novel hum
C 29	101	100.0	776	10	ADC74663 Human sec
C 30	101	100.0	776	10	ADA57709 BAC fragm
C 31	101	100.0	776	12	ADN41551 Novel hum
C 32	101	100.0	845	4	AAS30559 DNA encod
C 33	101	100.0	845	4	AAS27818 DNA encod
C 34	101	100.0	845	4	ABK42983 Genomic s
C 35	101	100.0	845	4	AAS41807 Genomic s
C 36	101	100.0	845	4	AAS41855 Genomic s
C 37	101	100.0	845	4	AAK85485 Human imm
C 38	101	100.0	845	4	AAK85434 Human imm
C 39	101	100.0	845	4	AAI07343 Human rep
C 40	101	100.0	845	4	AAI06587 Human rep
C 41	101	100.0	845	4	AAI07339 Human rep
C 42	101	100.0	845	4	AAI03228 Human rep
C 43	101	100.0	845	5	ABA14572 Human ner
C 44	101	100.0	845	5	AAS34680 Human DNA
C 45	101	100.0	845	9	ADB61139 Connectiv

ALIGNMENTS

RESULT 1  
AAV76919/c  
ID AAV76919 standard; DNA; 142 BP.  
XX  
AC AAV76919;  
DT 16-MAR-1999 (first entry)  
XX  
DB Staphylococcus aureus contig SEQ ID #2608.  
XX  
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.  
XX  
OS Staphylococcus aureus.  
XX  
FN EP786519-A2.  
XX  
PD 30-JUL-1997.  
XX  
PF 07-JAN-1997; 97EP-00100117.  
XX  
PR 05-JAN-1996; 96US-0009861P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;  
XX  
XX WPI; 1997-374922/35.  
XX  
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
XX stored on computer readable medium and used in the production of anti-  
XX S.aureus vaccines.  
XX  
PS Claim 1; Page 2287; 3271pp; English.  
XX  
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can  
 CC be used in a vaccine composition against *S.aureus* infection. The  
 CC polypeptides can also be used in a kit for the immunodetection of  
 CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,  
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
 CC for recombinant production of the polypeptides. The new DNA sequences  
 CC (and their fragments) are useful as primers or probes for isolating  
 CC homologues of any of the *S.aureus* DNA sequences contained on the computer  
 CC readable medium

SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 2.1e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATAGAAAAATAACAATAG 60

DB 107 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATAGAAAAATAACAATAG 48

QY 61 GGGTTCCGGCACATTTCCCGAAAAAGTCCACCTGACGTC 101

DB 47 GGGTTCCGGCACATTTCCCGAAAAAGTCCACCTGACGTC 7

#### RESULT 2

AAN10032/c

ID AAN10032 standard; DNA; 228 BP.

XX AC AAN10032;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;

XX KW insulin; ds.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT misc\_feature 1..4

FT /\*tag= a

FT /label= sticky end

FT 225..228

FT /\*tag= b

FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD ) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10039.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 3; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was  
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the  
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20  
 CC nucleotides before its translational start signal. In the examples, the  
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
 CC fragment (CB6) for rat preproinsulin (see AAN10034)

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 2.3e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATAGAAAAATAACAATAG 60

DB 175 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATAGAAAAATAACAATAG 116

QY 61 GGGTTCCGGCACATTTCCCGAAAAAGTCCACCTGACGTC 101

DB 115 GGGTTCCGGCACATTTCCCGAAAAAGTCCACCTGACGTC 75

#### RESULT 3

AAN10031/c

ID AAN10031 standard; DNA; 251 BP.

XX AC AAN10031;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;

XX KW insulin; ds.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT misc\_feature 1..4

FT /\*tag= a

FT /label= sticky end

FT 248..251

FT /\*tag= b

FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD ) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10038.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 2; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
 CC fragment (CB6) for rat preproinsulin (see AAN10034)  
 XX  
 SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGAATGTATTAGAAAAATAAACAATAG 60  
 Db |||||  
 175 AGGGTTATTGCTCATGCGCGATACATATTGAATGTATTAGAAAAATAAACAATAG 116  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101  
 Db |||||  
 115 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 75  
 |||||

RESULT 4  
 AAV31229/c  
 ID AAV31229 standard; DNA; 400 BP.  
 XX AC AAV31229;  
 XX  
 DT 01-OCT-1998 (first entry)  
 XX  
 DE E. coli J96 pathogenicity island contig #43.  
 XX  
 KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;  
 KW PAI V; phev; vaccine; protective immune response; ds.  
 XX  
 OS Escherichia coli.  
 XX  
 XX WO9822575-A2.  
 XX  
 XX 28-MAY-1998.  
 PD  
 XX  
 XX 21-NOV-1997; 97WO-US021347.  
 XX  
 XX 22-NOV-1996; 96US-0031626P.  
 PR  
 XX 14-OCT-1997; 97US-0061953P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (UYWI-) UNIV WISCONSIN.  
 XX  
 XX Dillon PJ, Choi GH, Welch RA;  
 PI  
 XX WPI; 1998-312461/27.  
 DR  
 XX  
 XX New isolated uropathogenic E. coli nucleotide sequences - used to develop  
 PT products for the detection of pathogenic E. coli and to elicit an immune  
 PT response to pathogenic E. coli.  
 XX  
 XX Claim 21; Page 140-141; 250pp; English.  
 PS  
 XX This sequence represents a E. coli strain J96 contig containing  
 CC pathogenicity island (PAI) sequences, and represents a nucleic acid  
 CC molecule of the invention. PAIs are large fragments of DNA which comprise  
 CC pathogenicity determinants. The sequences of the invention are taken from  
 CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near phev)  
 CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
 CC approximately 94 min (at pHER) on the E. coli chromosome and is  
 CC approximately 160 kb in size. Antibodies specific to the proteins encoded  
 CC by the PAI open reading frames of the invention can be used in kits to  
 CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
 CC a protective immune response in an animal to the uropathogenic E. coli  
 CC strain J96  
 XX  
 XX Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGAATGTATTAGAAAAATAAACAATAG 60  
 Db |||||  
 165 AGGGTTATTGCTCATGCGCGATACATATTGAATGTATTAGAAAAATAAACAATAG 106  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101  
 Db |||||  
 105 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 65  
 |||||

RESULT 5  
 AAN60624/c  
 ID AAN60624 standard; DNA; 456 BP.  
 XX AC AAN60624;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 29-OCT-1991 (first entry)  
 XX  
 DE Plasmid pUG201 sequence encoding beta-urogastrone.  
 XX  
 KW Beta-lactamase signal peptide; pGH54; pGH55; ss.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT promoter 125..170  
 FT /\*tag= a  
 FT RBS 200..203  
 FT /\*tag= b  
 FT CDS 209..439  
 FT /\*tag= c  
 FT sig\_peptide 209..277  
 FT /\*tag= d  
 FT /label= Beta-lactamase signal peptide  
 FT mat\_peptide 278..436  
 FT /\*tag= e  
 FT /label= Beta-urogastrone  
 XX  
 XX WO8603779-A.  
 XX  
 XX 03-JUL-1986.  
 PD  
 XX  
 XX 19-DEC-1985; 85WO-JP000696.  
 PF  
 XX  
 XX 21-DEC-1984; 84JP-00271206.  
 PR  
 XX  
 XX (EART ) EARTH CHEM CO LTD.  
 PA (OHGA/) OHGAI H.  
 XX  
 XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;  
 PI  
 XX WPI; 1986-182911/28.  
 DR P-PSDB; AAP60678.  
 XX  
 XX Recombinant vector for polypeptide secretion - contains signal peptide  
 PT sequence directly bonded to peptide-coding sequence.  
 PT  
 XX Disclosure; Table 4; 79pp; Japanese.  
 PS  
 XX The plasmid produces secreted beta-urogastrone in a transformed  
 CC expression system. Similar plasmids may be constructed where the  
 CC secretion signal may be coupled with eg. somatostatin, insulin, growth  
 CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,  
 CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60  
 |||  
 Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114  
 |||  
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
 |||  
 Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73  
 |||

## RESULT 6

AAAN71080/c  
 ID AAN71080 standard; DNA; 456 BP.  
 XX  
 AC AAN71080;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 13-MAY-1991 (first entry)  
 XX  
 XX Sequence encoding beta-urogastrone.  
 XX  
 XX pUGT 150s; beta-UG; ds.  
 KW  
 XX Escherichia coli.  
 OS Synthetic.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT promoter 125..170  
 FT /\*tag= a  
 FT CDS 209..439  
 FT /\*tag= b  
 FT /\*transl\_except= (pos:434..436,aa:Arg)  
 FT  
 XX JP62190083-A.

FN XX  
 XX 20-AUG-1987.  
 XX  
 XX 14-FEB-1986; 86JP-00031415.  
 XX  
 PR 14-FEB-1986; 86JP-00031415.  
 XX  
 XX (EART ) EARTH SEIYAKU KK.  
 PA  
 XX  
 XX WPI; 1987-273761/39.  
 DR  
 XX  
 XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.  
 PT  
 XX  
 XX Disclosure; Page 553; 34pp; Japanese.  
 PS  
 XX  
 XX Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing sequences comprising a tac promoter, SD site, signal peptide, and coding sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60  
 |||  
 Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114  
 |||  
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
 |||  
 Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73  
 |||

## RESULT 7

AAAN70833/c  
 ID AAN70833 standard; DNA; 456 BP.  
 XX  
 AC AAN70833;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 18-JAN-1991 (first entry)  
 XX  
 XX Beta-urogastrone sequence.  
 DE  
 XX Tumour; inosine; DNA probe; ds.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT promoter 125..170  
 FT /\*tag= b  
 FT RBS 200..204  
 FT /\*tag= c  
 FT CDS 209..439  
 FT /\*tag= a  
 FT sig\_peptide 209..277  
 FT /\*tag= d  
 FT  
 XX JP62244398-A.  
 PN  
 XX 24-OCT-1987.  
 PD  
 XX 16-APR-1986; 86JP-00087368.  
 PF  
 XX 16-APR-1986; 86JP-00087368.  
 PR  
 XX (SEKI ) SEKISUI CHEM IND CO LTD.  
 PA  
 XX  
 XX WPI; 1987-339045/48.  
 DR P-PSDB; AAP70505.  
 DR  
 XX  
 XX Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.  
 PT  
 XX  
 XX Disclosure; Page 11; 11pp; Japanese.  
 PS  
 XX  
 XX An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The ssDNA and probe are hybridized and the existence of DNA in the product is detected. It can be used to detect the presence of malignant tumour.  
 CC  
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60  
 |||  
 Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114  
 |||  
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101  
 |||  
 Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 73  
 |||

## RESULT 8

AAAN81765/c  
 ID AAN81765 standard; DNA; 456 BP.  
 XX  
 AC AAN81765;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 13-DEC-1990. (first entry)  
 DT

```
XX DE Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
XX DE Arg (53).
XX KW Gastric acid secretion; cell proliferation; hormone; ds.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT CDS 209..277
XX FT FT /*tag= a
XX FT FT 278..439
XX FT FT /*tag= b
XX FT FT /product= "New beta-urogastrone deriv."
XX PN JP63012298-A.
XX PD 19-JAN-1988.
XX PF 30-JUN-1986; 86JP-00153783.
XX PR 30-JUN-1986; 86JP-00153783.
XX PA (EART ) EARTH SEIYAKU KK.
XX DR WPI; 1988-054638/08.
XX DR P-PSDB; AAP81349.
XX PT New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX PT proliferation promotion activity.
XX PS Disclosure; Page 685; 76pp; Japanese.
XX CC The deriv. has various biological activities such as gastric acid
XX CC secretion inhibiting action, or cell proliferation promoting action. The
XX CC deriv. has the same biological or pharmacological activities as beta-
XX CC urogastrone. It is not susceptible to denaturation by oxidn. and is
XX CC chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX CC pepsinase. (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGTATTTAGAAAAATAACAAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGTATTTAGAAAAATAACAAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
RESULT 9
ABA90413/c
ID ABA90413 standard; DNA; 466 BP.
XX AC ABA90413;
XX DT 12-FEB-2002 (first entry)
XX DE Drosophila cell cycle progression protein coding sequence #48.
XX KW Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
XX KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
XX KW cell cycle progression protein; tumour; proliferative disorder;
XX KW cardiovascular; autoimmune; dermatological disorder; ds.
XX OS Drosophila sp.
XX PI Fraser CW;
XX
```

```
PN WO200172774-A2.
XX PD 04-OCT-2001.
XX PF 23-MAR-2001; 2001WO-GB001297.
XX PR 24-MAR-2000; 2000GB-00007268.
XX PA (CYCL-) CYCLACEL LTD.
XX PI Deak P, Glover DM, Midgley C;
XX FT WPI; 2002-055132/07.
XX DR Polynucleotides encoding cell cycle progression proteins, useful for
XX FT treating a tumor or a proliferative disorder.
XX PT Claim 1; Page 99; 213pp; English.
XX CC The present invention relates to Drosophila cell cycle progression
XX CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
XX CC ABA90520). The coding sequences and proteins are useful for identifying a
XX CC substance capable of affecting the function of the corresponding gene, a
XX CC substance capable of inhibiting the cell division cycle, or capable of
XX CC inhibiting mitosis and/or meiosis. They can also be used in a method for
XX CC treating a tumour or proliferative disorder, cardiovascular disorders
XX CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
XX CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
XX CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
XX CC disorders (such as malaria)
XX SQ Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
XX Query Match 100.0%; Score 101; DB 6; Length 466;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGTATTTAGAAAAATAACAAATAG 60
Db 280 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGTATTTAGAAAAATAACAAATAG 221
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 220 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180
RESULT 10
AAAX211173/c
ID AAAX211173 standard; DNA; 487 BP.
XX AC AAAX211173;
XX DT 05-MAY-1999 (first entry)
XX DE Polynucleotide sequence from the genome of Treponema pallidum.
XX DE Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX KW enzyme production; ds.
XX OS Treponema pallidum.
XX PN WO9859034-A2.
XX PD 30-DEC-1998.
XX PF 23-JUN-1998; 98WO-US013041.
XX PR 24-JUN-1997; 97US-0050667P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Fraser CW;
XX
```







PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234233P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249247P.  
PR 17-NOV-2000; 2000US-0249248P.  
PR 17-NOV-2000; 2000US-0249249P.  
PR 17-NOV-2000; 2000US-0249250P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-476223/51.

XX Novel isolated prostate gland related polypeptide useful for diagnosis  
PT and treatment of disorders of prostate such as prostatodystonia,  
PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.

PS Claim 1; SEQ ID NO 418; 512pp; English.

XX The invention relates to novel isolated prostate gland related nucleic  
CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,  
CC prognosis, prevention, and/or treatment of diseases and/or disorders of  
CC the prostate such as acute non-bacterial prostatitis, chronic non-  
CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,  
CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic  
CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including  
CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and  
CC squamous cell carcinomas. (I), (II) and antibody to (I) are useful for  
CC diagnosing and treating reproductive system disorders (Paget's disease),  
CC autoimmune disorders (systemic lupus erythematosus, rheumatoid  
CC arthritis), blood-related disorders (sickle cell anaemia),  
CC hyperproliferative disorders, urinary system disorders  
CC (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory  
CC disorders, musculoskeletal system disorders, neural activity and  
CC neurological disorders (Alzheimer's disease and Parkinson's disease),  
CC endocrine disorders (Addison's disease), gastrointestinal disorders  
CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),  
CC pancreatic and gall bladder disorders, disorders of the large intestine,  
CC developmental and inherited disorders, diseases at the cellular level,  
CC and wound healing and epithelial cell proliferation. (I) or (II) is  
CC useful to prevent skin aging, for preventing hair loss, to maintain  
CC organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;

Best Local Similarity 100.0%; Pred. No. 2.9e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGGAATGTTATTTAGAAAATAACAATAG 60  
|||||  
Db 546 AGGGTTATTGTCATGAGCGGATACATATTGGAATGTTATTTAGAAAATAACAATAG 487  
|||||

Qy 61 GGTTTCGGCGCACATTTCCCGAAAGTGCACCTGACGTC 101  
|||||

Db 486 GGTTTCGGCGCACATTTCCCGAAAGTGCACCTGACGTC 446  
|||||

RESULT 15

AAS27819/c

ID AAS27819 standard; DNA; 776 BP.

[illegible]

PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465460/50.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders.  
XX  
XX Claim 1; SEQ ID NO 1479; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative disorders  
CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
CC respiratory disorders, dermatological disorders (e.g. Addison's  
CC epithelial cell proliferation, endocrine disorders (e.g. Addison's  
CC disease), reproductive system disorders, gastrointestinal disorder  
CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
CC B-cell responsiveness to pathogens, activators of T-cells, to induce  
CC higher affinity antibodies, and as a means to induce tumour proliferation  
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-  
CC AAS27850 represent novel signal transduction pathway protein coding  
CC sequences and PCR primers of the invention  
XX  
SQ Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;  
Query Match 100.0%; Score 101; DB 4; Length 776;  
Best Local Similarity 100.0%; Pred. No. 2.9e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTTAGAAAAATAACAATAG 60  
Db |||||||TTCCTCATGAGCGGATACATATTGTAATGTTATTTAGAAAAATAACAATAG 487  
Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101  
Db |||||||TTCCTCATGAGCGGATACATATTGTAATGTTATTTAGAAAAATAACAATAG 487  
Qy 486 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 446  
Db |||||||TTCCTCATGAGCGGATACATATTGTAATGTTATTTAGAAAAATAACAATAG 487

Search completed: July 14, 2005, 07:01:51  
Job time : 142.448 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-65\_COPY\_8384\_8484

Perfect score: 101

Sequence: 1 aggtttattgtctatgagc.....gaaaagtgcacctgacgtc 101

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl1:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095 83374 Heb
C 2	101	100.0	300	5	BU963956 EST88 Heb
C 3	101	100.0	300	5	BU964094 EST226 He
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL597149 DKFZp313J
C 6	101	100.0	414	4	CC819240 100005D19
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819233 100006J13
C 9	101	100.0	495	4	BI805285 S035A01 S
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523 100004L13
C 12	101	100.0	503	9	CC819854 100006N08
C 13	101	100.0	515	9	CC817752 100003C16
C 14	101	100.0	518	9	CC817128 100002D21
C 15	101	100.0	519	9	CC817162 100002J19
C 16	101	100.0	519	9	CC817796 100003K14
C 17	101	100.0	521	9	CC819067 100005C09
C 18	101	100.0	533	9	CC819841 100006L07
C 19	101	100.0	542	9	CC816892 100002L01
C 20	101	100.0	550	7	CR766622 DKFZp469H
C 21	101	100.0	551	9	CC816905 100002N02
C 22	101	100.0	554	9	CC819058 100005A09
C 23	101	100.0	563	9	CC819270 100005G21
C 24	101	100.0	566	9	CC816848 100002D02

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

## ALIGNMENTS

RESULT 1  
BM078095/c  
LOCUS BM078095 300 bp mRNA linear EST 30-NOV-2001  
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma  
cylindrosporum cDNA 5', mRNA sequence.  
ACCESSION BM078095  
VERSION BM078095.1 GI:17157967  
KEYWORDS EST.  
SOURCE Hebeloma cylindrosporum  
ORGANISM Hebeloma cylindrosporum  
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Agaricales; Cortinariaceae; Hebeloma.  
REFERENCE 1 (bases 1 to 300)  
AUTHORS Wipf D., Benjdia M., Tegeder M. and Frommer W.B.  
TITLE Construction of a functional cDNA library from the ectomycorrhizal  
fungus Hebeloma cylindrosporum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wipf D  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: pDR196 5' primer (PNA 5')  
HIGH quality sequence stop: 300  
POLYA=NO.

## FEATURES

Location/Qualifiers  
1..300  
/organism="Hebeloma cylindrosporum"  
/mol\_type="mRNA"  
/strain="H1"  
/db\_xref="taxon:76867"  
/tissue\_type="Mycelia"  
/lab\_host="E. coli XL1-Blue"  
/clone\_lib="Hebeloma cylindrosporum functional cDNA  
library"  
/note="vector: pDR 196 (unpublished); Site\_1: EcoRI;  
Site\_2: XhoI"

## ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 8.1e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 AGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60

```

|||||
174 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST226 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 111
|||||

RESULT 3
BU964094/c
LOCUS
DEFINITION EST226 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU964094
VERSION BU964094.1 GI:24204891
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 111
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGTTATTGCTCATGCGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 111
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 39 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 98
QY 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 5
LOCUS
DEFINITION
AL597149 391 bp mRNA linear EST 04-SEP-2003
DFK2p313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
AL597149
VERSION
KEYWORDS
SOURCE
    Homo sapiens (human)
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 391)
    Koehler, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.
    EST (Koehler, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.)
    Unpublished (1999)
    Contact: MIPS
    MIPS
    Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
    This is the 5' sequence of the clone insert
    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
    sequenced by BMEP (Biomedical Research Center at the Charite,
    Berlin/Germany) within the cDNA sequencing consortium of the German
    Genome Project.
    No sl sequence available.
    This clone (DKF2p313J1611) is available at the RZPD in Berlin.
    Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
    Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKF2p313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 99 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 39 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 98
QY 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 6
LOCUS
DEFINITION
CC819240/c 414 bp DNA linear GSS 17-JUL-2003
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION
    CC819240
VERSION
    CC819240.1 GI:32899308
KEYWORDS
    GSS.
SOURCE
    Sterkiella histriomuscorum (Oxytricha trifallax)
    Sterkiella histriomuscorum
    Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
    Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
    Dunn, D., Doak, T., Herrick, G. and Weiss, R.
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
    Unpublished (2003)
    Contact: Robert B. Weiss
    University of Utah Genome Center
    University of Utah
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T4-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: FWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                pWD42 (GI|4732114|9b|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 414 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 355
QY 61 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 354 GGGTTCGGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 314
```

```

RESULT 7
BJ684174/c
LOCUS BJ684174 HCST library Haplochromis chilotes cDNA clone no90c12, 417 bp mRNA linear EST 23-APR-2004
DEFINITION BJ684174 mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Labroides; Cichlidae; Haplochromis.
1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
Orf sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES             source
    1..417
        /organism="Haplochromis chilotes"
        /mol_type="mRNA"
        /db_xref="taxon:257977"
        /clone="no90c12"
        /tissue_type="jaw"
        /dev_stage="varied"
        /clone_lib="HCST library"
ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAACAATAG 60
    |||||||
Db 129 AGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAACAATAG 70
    |||||||
Qy 61 GGTTCCGGCAGATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||||
Db 69 GGTTCCGGCAGATTTCCTCCGAAAAGTGCACCTGACGTC 29
    |||||||

RESULT 8
CC819923/c
LOCUS CC819923 491 bp DNA linear GSS 17-JUL-2003
DEFINITION 10006J13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC10006J13 R, genomic survey
sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
FEATURES             source
    1..495
        /organism="Oryza sativa"
        /mol_type="mRNA"
        /db_xref="taxon:4530"
        /clone="S035A01"

```

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES             source
    1..491
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC10006J13"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAACAATAG 60
    |||||||
Db 412 AGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAACAATAG 353
    |||||||
Qy 61 GGTTCCGGCAGATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||||
Db 352 GGTTCCGGCAGATTTCCTCCGAAAAGTGCACCTGACGTC 312
    |||||||

RESULT 9
BI805285
LOCUS BI805285 495 bp mRNA linear EST 02-OCT-2001
DEFINITION S035A01 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES             source
    1..495
        /organism="Oryza sativa"
        /mol_type="mRNA"
        /db_xref="taxon:4530"
        /clone="S035A01"

```



```

/tissue_type="Stem"
/dev stage="3-5 leaf stage"
/clone lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/notes="Vector: pSport2"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 62 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 121
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||
Db 122 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 162
    |||||

RESULT 10
CC818374/c
LOCUS
DEFINITION
100004807R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004B07 R, genomic survey
sequence.
ACCESSION
CC818374
VERSION
CC818374.1 GI:32897661
KEYWORDS
GSS.
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 495)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.

FEATURES
    source
    1..495
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC100004B07"
        /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /note="Vector: PWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        adapted vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 332
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||

```

```

Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 333
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||
Db 332 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 292
    |||||

RESULT 11
CC818523/c
LOCUS
DEFINITION
100004L13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION
CC818523
VERSION
CC818523.1 GI:32897943
KEYWORDS
GSS.
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 496)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.

FEATURES
    source
    1..496
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC100004L13"
        /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /note="Vector: PWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adapted mouse DNA was annealed to
        adapted vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 332
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||

```



Stichotrichida; Oxytrichidae; Sterkiella.  
 1 (Bases 1 to 518)  
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.  
 Paired end reads from plasmid inserts of Oxytricha trifallax  
 macronuclear chromosomes  
 Unpublished (2003)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dunn@genetics.utah.edu  
 Plate: 0002 row: D column: 21  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 518.

FEATURES  
 source  
 1..518  
 Location/Qualifiers  
 /organism="Sterkiella histriomuscorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:94289"  
 /clones="UUC100002D21"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Oxytricha plasmid UUGC10 library"  
 /note="Vector: PWD42nv; Purified macronuclear chromosomal  
 DNA from Oxytricha trifallax was blunt end-repaired with  
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
 oligonucleotides were ligated to the blunt ends in high  
 molar excess. Vector DNA was prepared from a derivative of  
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
 derivative of plasmid R1. The vector was ligated with  
 adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. Coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAACAATAG 60  
 Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAACAATAG 351

Oy 61 GGGTTCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101  
 Db 350 GGGTTCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 310

RESULT 15  
 CC817162/c  
 LOCUS  
 DEFINITION  
 100002J19 Oxytricha plasmid UUGC10 library Sterkiella  
 histriomuscorum genomic clone UUC100002J19 R, genomic survey  
 sequence.  
 CC817162  
 CC817162.1 GI:32896449  
 GSS.  
 SOURCE  
 ORGANISM  
 Sterkiella histriomuscorum (Oxytricha trifallax)  
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;  
 Stichotrichida; Oxytrichidae; Sterkiella.  
 1 (Bases 1 to 519)  
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.  
 Paired end reads from plasmid inserts of Oxytricha trifallax  
 macronuclear chromosomes  
 Unpublished (2003)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center

University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dunn@genetics.utah.edu  
 Plate: 0002 row: J column: 19  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 519.

FEATURES  
 source  
 1..519  
 Location/Qualifiers  
 /organism="Sterkiella histriomuscorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:94289"  
 /clone="UUC100002J19"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Oxytricha plasmid UUGC10 library"  
 /note="Vector: PWD42nv; Purified macronuclear chromosomal  
 DNA from Oxytricha trifallax was blunt end-repaired with  
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
 oligonucleotides were ligated to the blunt ends in high  
 molar excess. Vector DNA was prepared from a derivative of  
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
 derivative of plasmid R1. The vector was ligated with  
 adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. Coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAACAATAG 60  
 Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAACAATAG 357

Oy 61 GGGTTCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101  
 Db 356 GGGTTCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 316

Search completed: July 14, 2005, 23:23:29  
 Job time : 962.667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 gacggatcggagatctccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3986	12 PCDNA32EO	X90639 cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Expressio
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5082	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CV089673	U89673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A44171	A44171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222266	AR222266 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CV089672	U89672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1	AR098190	Sequence 5 from patent US 6074850.	3853 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR098190					
DEFINITION	AR098190					
ACCESSION	AR098190.1	GI:12807447				
VERSION	Unknown.					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3853)					
AUTHORS	Antelman,D., Gregory,R.J. and Wills,K.N.					
TITLE	Retinoblastoma fusion polypeptides					
JOURNAL	Patent: US 6074850-A 5 13-JUN-2000;					
FEATURES	Location/Qualifiers					
source	1..3853					
	/organism="unknown"					
	/mol_type="unassigned DNA"					

Query Match	100.0%;	Score 100;	DB 6;	Length 3853;
Best Local Similarity	100.0%;	Pred. No. 9.4e-24;		
Matches 100;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGCTCTCAGTACAACTCTGCTCTCATG	60	
Db	1	GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGCTCTCAGTACAACTCTGCTCTCATG	60	
Qy	61	CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT	100	
Db	61	CCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTGTT	100	
RESULT 2				
AR207832				
LOCUS	AR207832	3853 bp	DNA	linear
DEFINITION	Sequence 5 from patent US 6379927.			
ACCESSION	AR207832			
VERSION	AR207832.1	GI:21507688		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			

Unclassified.  
1 (bases 1 to 3853)  
Antelman,D., Gregory,R.J. and Wills,K.N.  
TITLE Retinoblastoma fusion proteins  
JOURNAL Patent: US 6379927-A 5 30-APR-2002;  
FEATURES Location/Qualifiers  
source  
1. 3853  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 3853;  
Best Local Similarity 100.0%; Pred. No. 9.4e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|  
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|

RESULT 3  
BD009729  
LOCUS 3853 bp DNA linear PAT 31-JAN-2002  
DEFINITION Tissue specific expression of retinoblastoma protein.  
ACCESSION BD009729  
VERSION BD009729.1 GI:18638102  
KEYWORDS JP 2001503638-A/3.  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

REFERENCE  
1 (bases 1 to 3853)  
Antelman,D., Gregory,R.J. and Wills,K.N.  
TITLE Tissue specific expression of retinoblastoma protein  
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;  
COMMENT CANJI INC  
OS Unidentified  
PN JP 2001503638-A/3  
PD 21-MAR-2001  
PF 13-NOV-1997 JP 1998522958  
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PT  
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC  
C07H21/04,C07K5/00,A61K38/00,A61K35/12  
CC Strandedness: Single;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT CDS 209..862.  
Location/Qualifiers  
1. 3853  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

FEATURES  
source  
1. 3853  
/organism="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 3853;  
Best Local Similarity 100.0%; Pred. No. 9.4e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|  
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|

RESULT 4  
PCDNA3ZEO  
LOCUS 3986 bp DNA linear SYN 16-AUG-1995  
DEFINITION Cloning vector pCDNA3ZEO DNA.  
ACCESSION X90639  
VERSION X90639.1 GI:949972  
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE  
1 Peters,H., Hunthausen,T., Kroenke,M. and Marget,M.  
AUTHORS A new small sized high-level eukaryotic expression vector  
TITLE Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 3986)  
Peters,H.  
AUTHORS Direct Submission  
TITLE Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,  
JOURNAL Michaelstr.5, D- 24105 Kiel, FRG  
COMMENT Related sequences: M21295 and K03104.  
FEATURES Location/Qualifiers  
source  
1. 3986  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
/plasmid="pCDNA3ZEO"  
1. 2125  
/note="Cloning vector (pCDNA3) (Invitrogen)"  
misc\_feature 889..994  
/note="multiple cloning site (MCS)"  
misc\_feature 2126..2796  
/note="cloning vector (pZeoSV) (Invitrogen)"  
misc\_feature 2797..3986  
/note="cloning vector (pCDNA3)"

ORIGIN  
Query Match 100.0%; Score 100; DB 12; Length 3986;  
Best Local Similarity 100.0%; Pred. No. 9.3e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|  
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTT 100  
|

RESULT 5  
AR098191  
LOCUS 4026 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 19 from patent US 6074850.  
ACCESSION AR098191  
VERSION AR098191.1 GI:12807448  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE  
1 (bases 1 to 4026)  
Antelman,D., Gregory,R.J. and Wills,K.N.  
AUTHORS Retinoblastoma fusion polypeptides  
TITLE Retinoblastoma fusion polypeptides  
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;  
FEATURES Location/Qualifiers  
source  
1. 4026  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 100; DB 6; Length 4026;  
Best Local Similarity 100.0%; Pred. No. 9.3e-24;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|  
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
|

```

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 6
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES Location/Qualifiers
    source 1..4026
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

RESULT 7
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN, RICHARD J GREGORY, KENNETH N WILLS PC
C07H21/04, C07K5/00, A61K38/00, A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4026
    /organism="Unidentified".
    Location/Qualifiers
        1..4026
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

FEATURES
source
    1..4026
        /organism="unidentified"
        /mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 8
AR207833
LOCUS AR207833 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 10
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

```

```

Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 8
AR098192
LOCUS AR098192 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 10
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
    source 1..4249
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249 /organism='Unidentified'.
FEATURES
source
1..4249
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES
source
1..4341
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source
1..4341
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1..4597
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32644'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 14
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES
source
1..4341
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 15
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1..4597
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32644'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGATGGTTCGACTTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
```



```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 14
AX133940
LOCUS      AX133940      4840 bp      DNA      linear      PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AX133940
VERSION    AX133940.1 GI:14139881
KEYWORDS   .
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.
REFERENCE   1
AUTHORS     Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE       Cell transfection
JOURNAL     Patent: WO 0119853-A 1 22-MAR-2001;
            THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES    Location/Qualifiers
             1..4840
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
             /note="This sequence is artificial and is based on well
             established commercially available vectors that are cited
             with their vendor within the patent application"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 15
BD238492
LOCUS      BD238492      5053 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
            of using the same.
ACCESSION  BD238492
VERSION     BD238492.1 GI:33048262
KEYWORDS    JP 2002520000-A/18.
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.
REFERENCE   1 (bases 1 to 5053)
AUTHORS     Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
            and Chesnut,K.W.
TITLE       Expression vectors for stimulating an immune response and methods
            of using the same
JOURNAL     Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904, 15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
PC A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
FT /organism="Artificial Sequence".

FEATURES
Source      Location/Qualifiers
            1..5053
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

Search completed: July 14, 2005, 14:03:20
Job time : 752.127 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_1\_100  
Perfect score: 100  
Sequence: 1 gacggatcggagatctccc.....ctgtccctcgtgtgtgtt 100

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	1506	12 ADMA1035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADMA1037	Adm41037 Cytomegal
4	100	100.0	2241	12 ADMA1034	Adm41034 Human nuc
5	100	100.0	2294	12 ADMA1036	Adm41036 Cytomegal
6	100	100.0	3853	2 AAV40006	Aav40006 Plasmid p
7	100	100.0	4026	2 AAV40007	Aav40007 Plasmid p
8	100	100.0	4249	2 AAV63466	Aav63466 Plasmid p
9	100	100.0	4341	2 AAG62391	Aag62391 Vector pv
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pv
11	100	100.0	4341	6 AEN83143	Adn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pep2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA sequ
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rd

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 ADE21866	Ade21866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AAZ89476	Aaz89476 Transgeni
33	100	100.0	5446	6 AAS18619	Aas18619 Renilla l
34	100	100.0	5446	6 ABL53540	AbL53540 Vector pc
35	100	100.0	5446	12 ADN36314	Adn36314 Plasmid p
36	100	100.0	5458	6 ABL58494	AbL58494 Recombina
37	100	100.0	5458	6 ABL58493	AbL58493 Recombina
38	100	100.0	5543	6 ABK88868	Abk88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ade83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	AbL58489 Recombina
42	100	100.0	5614	6 ABL58490	AbL58490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AAI66195	Aai66195 Human FSH
45	100	100.0	5651	6 ABK40237	Abk40237 DNA encod

ALIGNMENTS

RESULT 1

ADMA1035	
ID	ADMA1035 standard; DNA; 1506 BP.
AC	ADMA1035;
XX	
DT	17-JUN-2004 (first entry)
XX	
DE	Fungus nucleotide sequence SEQ ID NO:3.
XX	
KW	engrafting foreign replacement cell; implanting foreign replacement cell;
KW	growth; differentiation; drug development; vaccine development;
KW	tissue transplantation; human disease study; fungus; gene; ds.
OS	Unidentified.
XX	
FN	WO2004027029-A2.
XX	
PD	01-APR-2004.
XX	
PF	17-SEP-2003; 2003WO-US029251.
XX	
PR	19-SEP-2002; 2002US-0411790P.
XX	(XIME-) XIMEREX INC.
PA	Beschorner WE, Sosa CE, Thompson SC;
PI	
XX	WPI; 2004-295402/27.
DR	
XX	Engrafting foreign replacement cells within a fetal non-human mammal,
PT	useful in producing chimeric mammals, comprises selectively destroying
PT	native cells in a tissue of a fetal non-human mammal host.
XX	
PS	Disclosure; SEQ ID NO 3; 48pp; English.
XX	
CC	The present invention describes a method for engrafting foreign
CC	replacement cells within a foetal non-human mammal, which comprises
CC	selectively destroying native cells in a tissue of a foetal non-human
CC	mammal host, where the number of maternal cells of the same tissue is not
CC	substantially reduced, and implanting foreign replacement cells in the
CC	tissue of the fetal non-human mammal host, where the foreign replacement
CC	cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a  
 CC mammalian host, and for producing chimeric mammals that can be used to  
 CC develop new drugs and vaccine, factors, drugs and tissues for  
 CC transplantation, also useful to study human diseases. The present  
 CC sequence represents a nucleotide sequence given in the Sequence Listing  
 CC of the present invention but not mentioned further within the  
 CC specification.

XX  
 SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;  
 Best Local Similarity 100.0%; Pred. No. 4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 |  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 |  
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
 |  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
 |

RESULT 2  
 ADH11349  
 ID ADH11349 standard; DNA; 1600 BP.  
 XX  
 AC ADH11349;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Vertebrate UNC-53 protein homologue related nucleotide sequence.  
 XX  
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
 KW cell shape regulator; cell motility regulator; cell migration;  
 KW cell behaviour regulator; phenotype; signal transduction pathway;  
 KW signal transducing protein; signal integrator protein;  
 KW neuronal regeneration; revascularisation; wound healing;  
 KW chronic neurodegenerative disease; acute traumatic injury;  
 KW fibrotic disease; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9824810-A2.  
 XX  
 PD 11-JUN-1998.  
 XX  
 XX 03-DEC-1997; 97WO-EP006956.  
 XX  
 PR 04-DEC-1996; 96GB-00025283.  
 XX  
 XX (JANC ) JANSSEN PHARM NV.  
 XX  
 XX Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verbasselt P;  
 XX Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
 XX Geysen J, Bogaert TAOR;  
 XX  
 XX WPI; 1998-362411/31.  
 XX P-PSDB; ADH11350.  
 XX  
 XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.  
 XX promoting neuronal regeneration, treating chronic neuro-degenerative  
 XX diseases or acute traumatic injuries.  
 XX  
 XX Disclosure; Page 410-411; 479pp; English.

CC The present invention describes a vertebrate protein homologue of an UNC-  
 CC 53 protein or Caenorhabditis elegans or a functional equivalent,  
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)  
 CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of  
 CC cell shape, motility, or the direction of cell migration for use as a  
 CC therapeutic; (7) a method for determination of whether a protein is an  
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
 CC motility or the direction of migration by contacting a host cell  
 CC expressing a homologue of UNC-53 and determining a change of phenotype;  
 CC (8) a method for identification of vertebrate homologues of C. elegans  
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
 CC a DNA library; and (9) a method for identification of a protein which is  
 CC active in the signal transduction pathway of a cell of which a vertebrate  
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
 CC antibody/homologue complex; and (iii) analysing such a complex to  
 CC identify any non-antibody protein bound to the complex. UNC-53 is a  
 CC signal transducing or signal integrator protein involved in controlling  
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate  
 CC homologues of UNC-53 can be used to promote neuronal regeneration,  
 CC revascularisation or wound healing, to treat chronic neurodegenerative  
 CC diseases or acute traumatic injuries or fibrotic diseases. The present  
 CC sequence is used in the exemplification of the present invention.

XX  
 SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 |  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 |  
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
 |  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
 |

RESULT 3  
 ADH41037  
 ID ADH41037 standard; DNA; 1782 BP.  
 XX  
 AC ADH41037;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.  
 XX  
 KW engrafting foreign replacement cell; implanting foreign replacement cell;  
 KW growth; differentiation; drug development; vaccine development;  
 KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.  
 XX  
 OS Cytomegalovirus.  
 XX  
 PN WO2004027029-A2.  
 XX  
 PD 01-APR-2004.  
 XX  
 XX 17-SEP-2003; 2003WO-US029251.  
 XX  
 XX 19-SEP-2002; 2002US-0411790P.  
 XX  
 XX (XIME-) XIMEREX INC.  
 XX  
 XX Beschoner WE, Sosa CE, Thompson SC;  
 XX WPI; 2004-295402/27.  
 XX  
 XX Engrafting foreign replacement cells within a fetal non-human mammal,  
 XX useful in producing chimeric mammals, comprises selectively destroying  
 XX native cells in a tissue of a fetal non-human mammal host.

CC Disclosure; SEQ ID NO 5; 48pp; English.  
 CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises  
CC selectively destroying native cells in a tissue of a foetal non-human  
CC mammal host, where the number of maternal cells of the same tissue is not  
CC substantially reduced, and implanting foreign replacement cells in the  
CC tissue of the foetal non-human mammal host, where the foreign replacement  
CC cells replace destroyed cells of the tissue. The method is useful for  
CC facilitating growth and differentiation of foreign cells within a  
CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX  
SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 12; Length 1782;  
Best Local Similarity 100.0%; Pred. No. 4.2e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

## RESULT 4

ADM41034  
ID ADM41034 standard; DNA; 2241 BP.

XX  
AC ADM41034;

XX  
DT 17-JUN-2004 (first entry)

XX  
XX Human nucleotide sequence SEQ ID NO:2.

XX  
XX engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX tissue transplantation; human disease study; human; gene; ds.

XX  
OS Homo sapiens.

XX  
XX WO2004027029-A2.

XX  
XX 01-APR-2004.

XX  
XX 17-SEP-2003; 2003WO-US029251.

XX  
XX 19-SEP-2002; 2002US-0411790P.

XX  
XX (XIME-) XIMEREX INC.

XX  
XX Beschornier WE, Sosa CE, Thompson SC;

XX  
XX WPI; 2004-295402/27.

XX  
XX Engrafting foreign replacement cells within a foetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a foetal non-human mammal host.

XX  
XX Disclosure; SEQ ID NO 2; 48pp; English.

XX  
XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the foetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to  
CC develop new drugs and vaccine, factors, drugs and tissues for  
CC transplantation, also useful to study human diseases. The present  
CC sequence represents a nucleotide sequence given in the Sequence Listing  
CC of the present invention but not mentioned further within the  
CC specification.

XX  
SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;  
Best Local Similarity 100.0%; Pred. No. 4.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

## RESULT 5

ADM41036  
ID ADM41036 standard; DNA; 2294 BP.

XX  
AC ADM41036;

XX  
DT 17-JUN-2004 (first entry)

XX  
DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX  
XX engrafting foreign replacement cell; implanting foreign replacement cell;  
XX growth; differentiation; drug development; vaccine development;  
XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX  
OS Cytomegalovirus.

XX  
XX WO2004027029-A2.

XX  
XX 01-APR-2004.

XX  
XX 17-SEP-2003; 2003WO-US029251.

XX  
XX 19-SEP-2002; 2002US-0411790P.

XX  
XX (XIME-) XIMEREX INC.

XX  
XX Beschornier WE, Sosa CE, Thompson SC;

XX  
XX WPI; 2004-295402/27.

XX  
XX Engrafting foreign replacement cells within a foetal non-human mammal,  
XX useful in producing chimeric mammals, comprises selectively destroying  
XX native cells in a tissue of a foetal non-human mammal host.

XX  
XX Disclosure; SEQ ID NO 4; 48pp; English.

XX  
XX The present invention describes a method for engrafting foreign  
XX replacement cells within a foetal non-human mammal, which comprises  
XX selectively destroying native cells in a tissue of a foetal non-human  
XX mammal host, where the number of maternal cells of the same tissue is not  
XX substantially reduced, and implanting foreign replacement cells in the  
XX tissue of the foetal non-human mammal host, where the foreign replacement  
XX cells replace destroyed cells of the tissue. The method is useful for  
XX facilitating growth and differentiation of foreign cells within a  
XX mammalian host, and for producing chimeric mammals that can be used to  
XX develop new drugs and vaccine, factors, drugs and tissues for  
XX transplantation, also useful to study human diseases. The present  
XX sequence represents a nucleotide sequence given in the Sequence Listing  
XX of the present invention but not mentioned further within the  
XX specification.

```
SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 12; Length 2294;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 6
AAV40006
ID AAV40006 standard; DNA; 3853 BP.
AC AAV40006;
XX
XX
XX
XX 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
DE Plasmid pCTM.
XX
XX E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
XX Human cytomegalovirus.
OS
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FT promoter 209..864
FT /*tag= a
FT /*note= "CMV promoter"
FT misc_feature 907..1131
FT /*tag= b
FT /*function= "tripartite leader sequence"
FT promoter 1132..1149
FT /*tag= c
FT /*note= "SP6 promoter"
FT misc_feature 1679..3853
FT /*tag= d
FT /*note= "pUC19 backbone H3 to AatII"
FT CDS complement(2857..3717)
FT /*tag= e
FT /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
```

```
PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 4; 91pp; English.
XX
XX This is the nucleotide sequence of pCTM, a plasmid which contains a CMV
CC promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters,
CC and a multiple cloning site with a bovine growth hormone polyA site and
CC downstream SV40 polyA site. It has been used as a vector for the
CC expression of fusion proteins of the invention that comprise
CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
CC (see AAW62464). Such fusion proteins, particularly expressed from gene
CC therapy vectors, are used to treat hyperproliferative conditions,
CC specifically cancer (particularly of the bladder) or restenosis. They are
CC more effective in repressing transcription of the E2F promoter than RB
CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
XX SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 2; Length 3853;
Best Local Similarity 100.0%; Pred. No. 5.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 7
AAV40007
ID AAV40007 standard; DNA; 4026 BP.
XX
XX AAV40007;
XX
XX 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
DE Plasmid pCTMI.
XX
XX E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTMI; ss.
XX
XX Human cytomegalovirus.
OS
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FT promoter 209..864
FT /*tag= a
FT /*note= "CMV promoter"
FT misc_feature 907..1074
FT /*tag= b
FT /*function= "tripartite leader sequence"
FT intron 1075..1253
FT /*tag= c
FT /*note= "hybrid SV40 late intron"
FT promoter 1305..1322
FT /*tag= d
FT /*note= "SP6 promoter"
FT misc_feature 1851..4026
FT /*tag= e
FT /*note= "pUC19 backbone H3 to AatII"
```

```

FT CDS complement (3032..3890)
FT /*tag= f
FT /*note= "AMP-ORF"
PN WO9821228-A1.
PD 22-MAY-1998.
XX 13-NOV-1997; 97WO-US021821.
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX Antelman D, Gregory RJ, Wills KN;
XX WPI; 1998-297858/26.
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX Example 1; Fig 6; 91pp; English.
XX This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX vector. Plasmid pCTMI has been used as a vector for the expression of
XX fusion proteins of the invention that comprise retinoblastoma protein
XX (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX fusion proteins, particularly expressed from gene therapy vectors, are
XX used to treat hyperproliferative conditions, specifically cancer
XX (particularly of the bladder) or restenosis. They are more effective in
XX repressing transcription of the E2F promoter than RB alone and cause cell
XX -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX OS field.)
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 2; Length 4026;
Best Local Similarity 100.0%; Pred. No. 5.3e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
RESULT 8
AAV63466
ID AAV63466 standard; DNA; 4249 BP.
XX
XX AC AAV63466;
XX
XX 27-AUG-2003 (revised)
XX 15-FEB-1999 (first entry)
XX Plasmid pCTMIE.
XX E2F; transcription factor; human; retinoblastoma protein RB;
XX bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX thyroid hyperplasia; Grave's disease; psoriasis;
XX benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX Human cytomegalovirus.
XX OS mastadenovirus.
XX OS unidentified bacteriophage; T7.
XX OS unidentified bacteriophage; SP6.

```

---

```

OS OS Macaca mulatta; polyoma virus.
OS OS Bos taurus.
XX XX Chimeric.
XX Key Location/Qualifiers
XX promoter 209..864
XX /*tag= a
XX /*note= "CMV promoter"
XX misc_feature 907..1074
XX /*tag= b
XX /*function= "tripartite leader sequence"
XX intron 1081..1145
XX /*tag= c
XX /*note= "hybrid SV40 late intron"
XX mRNA 1164..1366
XX /*tag= d
XX /*note= "early mRNA"
XX enhancer 1261..1332
XX /*tag= e
XX /*note= "72 bp tandem repeat enhancer"
XX enhancer 1333..1404
XX /*tag= f
XX /*note= "72 bp tandem repeat enhancer"
XX misc_binding 1366
XX /*tag= g
XX /*note= "T antigen binding site"
XX intron 1372..1478
XX /*tag= h
XX /*note= "hybrid SV40 late intron"
XX promoter 1530..1545
XX /*tag= i
XX /*note= "SP6 promoter"
XX misc_feature 2075..4249
XX /*tag= j
XX /*note= "pUC19 backbone H3 to AatII"
XX CDS complement(3255..4113)
XX /*tag= k
XX /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX Example 1; Fig 8; 91pp; English.
XX This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX amplified product with BglII and inserting into BamHI-digested plasmid
XX pCMTI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX expression of fusion proteins of the invention that comprise
XX retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX (see AAW62464). Such fusion proteins, particularly expressed from gene
XX therapy vectors, are used to treat hyperproliferative conditions,
XX specifically cancer (particularly of the bladder) or restenosis. They are
XX more effective in repressing transcription of the E2F promoter than RB
XX alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 100; DB 2; Length 4249;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9  
 AAQ62391  
 ID AAQ62391 standard; DNA; 4341 BP.  
 XX  
 AC AAQ62391;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 18-NOV-1994 (first entry)  
 XX  
 DE Vector pVAC1.  
 XX  
 KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;  
 KW fusion protein; pSfi/Not.Tag1; pElB leader; human; immunoglobulin; VHL;  
 KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;  
 KW vaccine; ss.  
 XX  
 OS Synthetic.  
 XX

FH Key Location/Qualifiers  
 FT misc\_RNA complement(1. .775)  
 FT /\*tag= c  
 FT /\*note= "Claim 9"  
 FT 606. .780  
 FT /\*tag= b  
 FT /\*note= "Claim 8"  
 FT 606. .716  
 FT /\*tag= a  
 FT /\*note= "Claim 7"  
 XX

PN WO9408008-A1.  
 XX  
 PD 14-APR-1994.  
 XX  
 PF 04-OCT-1993; 93WO-GB002054.  
 XX  
 PR 02-OCT-1992; 92GB-00020808.  
 XX  
 PA (MEDI-) MEDICAL RES COUNCIL.  
 XX  
 PI Hawkins RE, Russell SJ, Stevenson FK, Winter GP;  
 XX  
 DR WPI; 1994-135575/16.  
 XX

PT Modulating immune response to a disease marker - by administering a  
 PT vector which expresses the disease marker to interact with the immune  
 PT system.  
 XX  
 PS Claim 10; Fig 7; 77pp; English.  
 XX

CC This sequence represents the vector pVAC1. This vector is based on the  
 CC commercially available vector pRC/RSV. Leader sequences and termination  
 CC signals were introduced into the vector to allow for production of fusion  
 CC proteins. The vector, pSfi/Not.Tag1, was modified to replace the pElB  
 CC leader with the human immunoglobulin VH1 leader sequence that permits the  
 CC encoding of an Sfil cloning site without modification of the amino acid  
 CC sequence. This fragment was then cloned as an EcoRI/Blunt-HindIII  
 CC fragment into NotI/Blunt- HindIII cut vector pRC/RSV to give pVAC1. The  
 CC single chain Fv for an individual patient can be inserted within the VH1  
 CC leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid  
 CC vaccine and it induces a strong humoral response to the antibody moiety  
 CC in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)  
 XX

SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 100; DB 2; Length 4341;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 10  
 AAS17704  
 ID AAS17704 standard; DNA; 4341 BP.  
 XX  
 AC AAS17704;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Vector pVAC1 encoding a DNA vaccine.  
 XX  
 KW Cytostatic; vaccine; tetanus toxin; FrC; tumour; CTL; PCR primer; pVAC1;  
 KW ds.  
 XX  
 OS Clostridium tetani.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Cauliflower mosaic virus.  
 XX  
 PN WO200179510-A1.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 17-APR-2001; 2001WO-GB001719.  
 XX  
 PR 17-APR-2000; 2000GB-00009470.  
 XX  
 PA (CANC-) CANCER RES VENTURES LTD.  
 XX  
 PI Rice J, Stevenson F;  
 XX  
 DR WPI; 2002-066370/09.  
 XX

PT Nucleic acid construct, useful to immunize against various diseases  
 PT including cancer, expresses the first domain of tetanus toxin FrC fused  
 PT to a disease peptide antigen to provide a vaccine.  
 XX  
 PS Disclosure; Fig 4; 71pp; English.  
 XX

CC The invention relates to a nucleic acid construct for delivery into  
 CC living cells in vivo, to induce an immune response to a disease peptide  
 CC antigen, where the construct directs expression of a fusion protein  
 CC comprising the peptide antigen and the first domain of FrC. Also included  
 CC are a nucleic acid vector comprising the above construct, a host cell  
 CC comprising the above construct or vector and a method of producing a  
 CC nucleic acid construct for inducing an immune response. The method  
 CC comprises identifying a nucleic acid sequence encoding a disease peptide  
 CC antigen comprising epitopes characteristic of the disease, cloning the  
 CC nucleic acid sequence, introducing the cloned nucleic acid into a vector  
 CC which allows the antigen to be expressed as a fusion with a first domain  
 CC FrC from tetanus toxin, and optionally isolating the construct from the  
 CC vector. The construct or vector is used as a vaccine to induce an immune  
 CC response, particularly to tumour antigens. The present sequence is vector  
 CC pVAC1 which encodes a vaccine of the invention  
 XX



SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 11  
ID ABN83143 standard; DNA; 4341 BP.  
AC ABN83143;  
DT 10-SEP-2002 (first entry)  
DE Plasmid pVAC1 complete sequence.  
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;  
KW cancer; B cell malignancy; ds.  
OS Synthetic.  
XX WO200240513-A2.  
XX 23-MAY-2002.  
XX 20-NOV-2001; 2001WO-GB005142.  
XX 20-NOV-2000; 2000GB-00028319.  
XX (CANC-) CANCER RES VENTURES LTD.  
XX Savelyeva N, Stevenson F;  
XX WPI; 2002-500202/53.  
XX Nucleic acid construct for delivery into living cells as a vaccine,  
XX useful for treating e.g. cancer, directs the expression of a fusion  
XX protein comprising an antigen and an adjuvant sequence derived from a  
XX plant viral coat protein.  
XX Example 3; Fig 7; 84pp; English.  
XX The invention relates to a novel nucleic acid construct for inducing an  
XX immune response in vivo to an antigen, capable of directing the  
XX expression of a fusion protein that comprises an antigen and an adjuvant  
XX sequence derived from a plant viral coat protein. The construct of the  
XX invention has cytostatic and virucide activity. The nucleic acid  
XX construct is useful for inducing an immune response in a patient, for  
XX vaccinating a patient against an infectious disease caused by an antigen  
XX derived from a pathogen e.g. a virus, for treating a cancer patient or a  
XX patient with a predisposition to cancer and for treating a patient having  
XX a B cell malignancy, where the construct is encapsidated, and optionally,  
XX a second nucleic acid sequence encoding a further immunomodulatory  
XX polypeptide is administered to the patient. The construct is also useful  
XX in medical treatment, and in the preparation of a vaccine for treating or  
XX preventing a disease state associated with the antigen. The sequence  
XX shows the complete sequence of vector pVAC1

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 5.4e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 12  
ID AAF24901 standard; DNA; 4597 BP.  
XX AAF24901;  
XX 20-APR-2001 (first entry)  
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.  
XX Microsphere; dihydrazide; hyaluronic acid; inflammatory response;  
KW myocardial ischemia; cardiac angiogenesis; haemophilia;  
KW vascular endothelial growth factor; VEGF; ss.  
XX Synthetic.  
XX WO2000078358-A2.  
XX PN FN  
XX 28-DEC-2000.  
XX 19-JUN-2000; 2000WO-US016937.  
XX 18-JUN-1999; 99US-0140260P.  
XX (COLL-) COLLABORATIVE GROUP LTD.  
XX Chen W;  
XX WPI; 2001-071363/08.  
XX Hyaluronic acid micro spheres for use in gene therapy of myocardial  
XX ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic  
XX acids crosslinked to nucleic acids.  
XX Example 1; Page 36-38; 38pp; English.  
XX The specification describes a microsphere comprising dihydrazide  
XX derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The  
XX microspheres cause reduced inflammatory responses, and have increased  
XX safety and biodegradability. The microspheres are useful for transfecting  
XX a cell of a subject and for treating a subject having myocardial  
XX ischemia, by increasing cardiac angiogenesis. They are also useful for  
XX treating haemophilia. The present sequence represents the plasmid  
XX pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is  
XX crosslinked to hyaluronic acid. The polynucleotide sequence encodes a  
XX vascular endothelial growth factor (VEGF)  
XX Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;  
XX Query Match 100.0%; Score 100; DB 4; Length 4597;  
XX Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 13  
AAD39652

AD39652 standard; DNA; 4639 BP.  
AD39652;  
22-OCT-2002 (first entry)  
Human small nuclear RNA (snRNA) DNA.  
Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;  
transgenic animal; ds.  
Homo sapiens.  
US2002058287-A1.  
16-MAY-2002.  
12-MAR-2001; 2001US-00804481.  
10-MAR-2000; 2000US-0188304P.  
(WHED) WHITEHEAD INST BIOMEDICAL RES.  
Graaf DD, Lander ES;  
WPI; 2002-499510/53.  
New recombinant vector containing sequence for small nuclear RNA, useful  
e.g. for identifying variant snRNA that suppresses expression of  
transcription products.  
Disclosure; Fig 1; 18pp; English.  
The invention relates to a recombinant vector which comprises DNA,  
consisting of an insertion cassette contained between at least two  
insertion sites, that encodes a small nuclear (sn) RNA. The invention is  
used to identify snRNA modifications that inhibit expression of  
transcription products (and the identified snRNA are used to suppress  
expression) for delivering antisense sequences to the nucleus and to  
create transgenic animals. The present DNA sequence is human snRNA, U1  
Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4639;  
Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
RESULT 14  
AAF83146  
ID AAF83146 standard; DNA; 4840 BP.  
AC AAF83146;  
XX AAF83146;  
DT 09-JUL-2001 (first entry)  
DE Complete sequence of vector pIRES/BS.  
KW Blastcidin resistance; BS gene; gene therapy; tissue engineering;  
cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;  
pIRES/BS; ss.  
OS Synthetic.  
XX WO200119853-A2.  
PN

AD39652 standard; DNA; 4639 BP.  
AD39652;  
22-OCT-2002 (first entry)  
Human small nuclear RNA (snRNA) DNA.  
Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;  
transgenic animal; ds.  
Homo sapiens.  
US2002058287-A1.  
16-MAY-2002.  
12-MAR-2001; 2001US-00804481.  
10-MAR-2000; 2000US-0188304P.  
(WHED) WHITEHEAD INST BIOMEDICAL RES.  
Graaf DD, Lander ES;  
WPI; 2002-499510/53.  
New recombinant vector containing sequence for small nuclear RNA, useful  
e.g. for identifying variant snRNA that suppresses expression of  
transcription products.  
Disclosure; Fig 1; 18pp; English.  
The invention relates to a recombinant vector which comprises DNA,  
consisting of an insertion cassette contained between at least two  
insertion sites, that encodes a small nuclear (sn) RNA. The invention is  
used to identify snRNA modifications that inhibit expression of  
transcription products (and the identified snRNA are used to suppress  
expression) for delivering antisense sequences to the nucleus and to  
create transgenic animals. The present DNA sequence is human snRNA, U1  
Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 6; Length 4639;  
Best Local Similarity 100.0%; Pred. No. 5.5e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
RESULT 14  
AAF83146  
ID AAF83146 standard; DNA; 4840 BP.  
AC AAF83146;  
XX AAF83146;  
DT 09-JUL-2001 (first entry)  
DE Complete sequence of vector pIRES/BS.  
KW Blastcidin resistance; BS gene; gene therapy; tissue engineering;  
cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;  
pIRES/BS; ss.  
OS Synthetic.  
XX WO200119853-A2.  
PN

XX 22-MAR-2001.  
XX 11-SEP-2000; 2000WO-GB003462.  
XX 11-SEP-1999; 99GB-00021418.  
XX (UYSH-) UNIV SHEPFIELD.  
XX Hollander AP, Barker MD, Kafienah W;  
XX WPI; 2001-290354/30.  
XX Novel nucleic acid molecule useful for therapeutic and cosmetic tissue  
XX engineering, comprising at least a functional part of blastcidin  
XX resistance gene linked through a recognition sequence, to a selected  
XX gene.  
XX Claim 11; Fig C; 44pp; English.  
XX The invention provides a nucleic acid molecule (I) comprising at least  
XX the functional part of blastcidin resistance (BS) gene, or its homolog,  
XX linked through a recognition sequence to at least one selected gene. (I)  
XX is useful in treatment comprising: (1) providing cells/tissues transfected  
XX with (I); (2) surgical administration of the cells/tissues to the patient  
XX to be treated; and optionally (3) monitoring the status of the cells/  
XX tissues by the patient. Therapeutic compositions comprising cells/tissues  
XX transformed with (I) is useful in identifying the role of genes in  
XX healthy and diseased tissue, in tissue engineering and in cosmetic  
XX surgery. Tissue engineering can be used to treat arthritis, joint  
XX replacement, skin grafts for burn victims, and replacement coronary  
XX arteries. Cosmetic tissue surgery includes rhinoplasty. The present  
XX sequence represents the nucleotide sequence of the vector pIRES/BS  
XX containing the BS gene  
SQ Sequence 4840 BP; 1154 A; 1227 C; 1236 G; 1223 T; 0 U; 0 Other;  
Query Match 100.0%; Score 100; DB 4; Length 4840;  
Best Local Similarity 100.0%; Pred. No. 5.6e-26;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60  
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100  
RESULT 15  
ADB33528  
ID ADB33528 standard; DNA; 5015 BP.  
XX ADB33528;  
XX ADB33528;  
XX 04-DEC-2003 (first entry)  
XX Expression vector nucleotide sequence SEQ ID NO:27.  
DE fusion protein; amyloid precursor protein; APP; transcription factor;  
KW neurotrophic; neuroprotective; APP inhibitor;  
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;  
KW gamma-secretase; human; gene; ds.  
XX Synthetic.  
OS Homo sapiens.  
XX WO2003072041-A2.  
PN 04-SEP-2003.  
XX 23-FEB-2003; 2003WO-US0005458.  
XX

Search completed: July 14, 2005, 07:01:24  
Job time : 147.038 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_1\_100

Perfect score: 100

Sequence: 1 gacggatcgagatctccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_hic.\*
- 4: gb\_est3.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_gsl1.\*
- 9: gb\_gsl2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	60.0	602	8	B67169 CpG0047A Cp
2	55.6	55.6	694	8	B2052929 jnr13903
3	55.6	55.6	696	8	B2050328 jnr42c12
4	55.6	55.6	717	8	B2054067 jnr38b09
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK119397 212c09.pl
9	53.4	53.4	766	7	CK120360 207j04.pl
10	53.4	53.4	788	7	CK117844 209p08.pl
11	53.4	53.4	898	9	CL141237 ISB1-118J
12	53.4	53.4	899	9	CL140877 ISB1-118B
13	53.4	53.4	1009	9	CL123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbe0049M
15	53	53.0	675	8	B2051815 jnr57d03
16	53	53.0	679	8	B2052857 jnr13903
17	53	53.0	700	8	B2050646 jnr66f08
18	53	53.0	701	8	B2052015 jnr56b03
19	53	53.0	708	8	B2054793 jnr33903
20	53	53.0	709	8	B2053587 jnr98d01
21	53	53.0	712	8	B2054005 jnr38b09
22	52.8	52.8	451	8	AQ863966 nbe0022E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	B2049113 jnr21d02

ALIGNMENTS

RESULT 1  
LOCUS B67169  
DEFINITION CpG0047A CpIOWAgDNA2 Cryptosporidium parvum genomic, linear GSS 12-MAY-2000  
sequence.  
ACCESSION B67169  
VERSION B67169.1 GI:2642750  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum  
ORGANISM Cryptosporidium parvum  
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
Cryptosporidiidae; Cryptosporidium.  
REFERENCE  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
MEDLINE 20183851  
PUBMED 10717299  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malariad@itsa.ucsf.edu  
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.  
Seq primer: T7  
Class: shotgun  
High quality sequence stop: 602.

FEATURES  
source

1..602  
Location/Qualifiers  
/organism="Cryptosporidium parvum"  
/mol\_type="genomic DNA"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/lab\_host="E. coli XL2 Blue MRF"  
/clone\_lib="CpIOWAgDNA2"  
/notes="Vector: PCR-Script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

# ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100  
 |||||  
 Db 1 CAGTACAAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 60

RESULT 2  
 BZ052929/c  
 LOCUS jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 DEFINITION sequence.

ACCESSION BZ052929.1 GI:23654922  
 VERSION BZ052929.1  
 KEYWORDS GSS.

ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 694)  
 AUTHORS Delehaanty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,  
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.

TITLE Whole genome shotgun reads from Brassica oleracea  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Plate: jnr13 row: 9 column: 03  
 Seq primer: -28RppOT reverse  
 Class: shotgun  
 High quality sequence start: 32  
 High quality sequence stop: 551.

FEATURES  
 source Location/Qualifiers  
 1..694  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOTw13; Whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea T01000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin. Genomic  
 DNA was provided by Pablo Rabinowicz (CSHL) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."

Query Match 55.6%; Score 55.6; DB 8; Length 694;  
 Best Local Similarity 77.9%; Pred. No. 9e-09;  
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGATGCC 62  
 |||||  
 Db 324 CGGATCGATAGGTCCTCGGAGTAGTTATGTCGACTCTCAGTACAAATCTGCTGATGCC 265  
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||

ORIGIN  
 Query Match 55.6%; Score 55.6; DB 8; Length 694;  
 Best Local Similarity 77.9%; Pred. No. 9e-09;  
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGATGCC 62  
 |||||  
 Db 324 CGGATCGATAGGTCCTCGGAGTAGTTATGTCGACTCTCAGTACAAATCTGCTGATGCC 265  
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||

ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 717)  
 AUTHORS Delehaanty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

RESULT 3  
 BZ050328  
 LOCUS jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 DEFINITION sequence.

ACCESSION BZ050328  
 VERSION BZ050328.1 GI:23649718  
 KEYWORDS GSS.

ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 696)  
 AUTHORS Delehaanty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,  
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.

TITLE Whole genome shotgun reads from Brassica oleracea  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Plate: jnr42 row: 9 column: 12  
 Seq primer: -21UPpOT forward  
 Class: shotgun  
 High quality sequence start: 35  
 High quality sequence stop: 180.

FEATURES  
 source Location/Qualifiers  
 1..696  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOTw13; Whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea T01000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin. Genomic  
 DNA was provided by Pablo Rabinowicz (CSHL) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."

Query Match 55.6%; Score 55.6; DB 8; Length 696;  
 Best Local Similarity 77.9%; Pred. No. 9e-09;  
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGATGCC 62  
 |||||  
 Db 45 CGGATCGATAGGTCCTCGGAGTAGTTATGTCGACTCTCAGTACAAATCTGCTGATGCC 104  
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||  
 Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

RESULT 4  
 BZ054067/c  
 LOCUS jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 DEFINITION sequence.

ACCESSION BZ054067  
 VERSION BZ054067.1 GI:23657216  
 KEYWORDS GSS.

ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 717)  
 AUTHORS Delehaanty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

TITLE  
JOURNAL  
COMMENT

Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
Whole genome shotgun reads from Brassica oleracea  
Unpublished (2002)  
Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Plate: jnr38 row: b column: 09  
Seq primer: -28RppOT reverse  
Class: shotgun  
High quality sequence start: 87  
High quality sequence stop: 543.

Location/Qualifiers

1. .717  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea001"  
/note="Vector: pOTW13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using Brassica oleracea T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

## FEATURES

source

1. .717  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea001"  
/note="Vector: pOTW13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using Brassica oleracea T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

## ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;  
Best Local Similarity 77.9%; Pred. No. 9.1e-09;  
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62  
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 188 GCATAGTTAAGCCAGCCCGACACC 163

## RESULT 5

AW409112

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW409112  
sal10h5 Salivary Gland Library Homo sapiens cdna, mRNA  
EST.  
AW409112.1 GI:11999687  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 348)  
Berghelm, A., Ogilvie, E., Arndt, S., Napier, H., Taylor, M., Lovett, M.,  
Simmons, A., Hide, M., and Ramsay, M.  
A high density transcript map between markers D8S550 and D8S1759 on  
9p22-23, using cDNA direct selection  
Unpublished (2000)  
Contact: Ramsay M  
Department of Human Genetics  
South African Institute For Medical Research  
P.O.Box 1038, Johannesburg, Gauteng, 2000, South Africa  
Fax: 2711 489 9226  
Email: micheler@mail.saimr.wits.ac.za.

Location/Qualifiers

1. .348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue type="Salivary Gland"  
/clone\_lib="Salivary Gland Library"  
/note="Vector: pAMP10"

## ORIGIN

Query Match 53.6%; Score 53.6; DB 2; Length 348;  
Best Local Similarity 80.5%; Pred. No. 4.1e-08;  
Matches 62; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 11 GAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTT 70  
Db 65 GCGTATACACCGCATATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTT 124

Qy 71 AAGCCAGTATCTGCTCC 87

Db 125 AAGCCAGTATACACTCC 141

## RESULT 6

AL715724/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

1. .343  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="BN0AA018ZF12"  
/tissue type="inner ear"  
/dev stage="embryonic"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cdna"  
/note="subtracted cdna library"

Location/Qualifiers

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 345)
AUTHORS Coimbra, R., Wei, D., Brottier, P., Blanchard, S., Levi, M.,
Hardelin, J.P., Weissenbach, J. and Petit, C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..345
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA0072C02"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cDNA"
/note="subtracted cDNA library"
ORIGIN
Query Match 53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76
DB 280 TTACACCGCATATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 221
QY 77 GTATCTGCTCC 87
DB 220 GTATACACTCC 210
RESULT 8
CKL19397/c
LOCUS CKL19397 761 bp mRNA linear EST 01-JUN-2004
DEFINITION 212009.p1 AtM1 Arabidopsis thaliana cDNA clone MPMGp2011009212
5-PRIME, mRNA sequence.
ACCESSION CKL19397
VERSION CKL19397.1 GI:47829713
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 761)
AUTHORS Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.
TITLE Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73, D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 761 Std Error: 0.00
Plate: 212 row: 0 column: 9
Seq primer: PQB65.
Location/Qualifiers
1..761
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:954234"
/db_xref="taxon:3702"
FEATURES
source
1..761
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:954234"
/db_xref="taxon:3702"
/clone="MPMGp2011J04207"
/tissue_type="inflorescence meristem"
/dev_stage="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="Vector: PQE-30NASt-attB (AY386205); Site_1: SalI;
/clone="MPMGp2011009212"
/tissue_type="inflorescence meristem"
/dev_stage="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="Vector: PQE-30NASt-attB (AY386205); Site_1: SalI;

```



expression clones. Average insert size is 1 kb. Note: The rearranged sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

Db 514 TTCAACCGCATATGGTGCACTCTCAGTACAATCTGTCTGATGCCGATAGTTAAGCA 455

QY 77 GTATCTGCTCC 87

Db 454 GTATACACTCC 444

```

LOCUS      CL140877              899 bp    DNA        linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
            genomic survey sequence.
ACCESSION  CL140877
VERSION    CL140877.1  GI:40634512
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 899)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   Location/Qualifiers
            source             1..899
                                /organism="Xenopus tropicalis"
                                /mol_type="genomic DNA"
                                /db_xref="taxon:8364"
                                /clone="ISB1-118B12"
                                /clone_lib="ISB1"
                                /note="vector: phelobAC11; ISB-1 Xenopus tropicalis BAC
                                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 899;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 76
Db 195 TTACACCGCATATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 136
QY 77 GTATCTGCTCC 87
Db 135 GTATACACTCC 125

RESULT 13
LOCUS      CL123953/c
DEFINITION ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
            genomic survey sequence.
ACCESSION  CL123953
VERSION    CL123953.1  GI:40617598
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1009)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   Location/Qualifiers
            source             1..1009
                                /organism="Xenopus tropicalis"
                                /mol_type="genomic DNA"
                                /db_xref="taxon:8364"
                                /clone="ISB1-84J15"
                                /clone_lib="ISB1"
                                /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 76
Db 252 TTACACCGCATATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 193
QY 77 GTATCTGCTCC 87
Db 192 GTATACACTCC 182

RESULT 14
LOCUS      AQ914559
DEFINITION nbe0049M21r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica
            cultivar-group) genomic clone nbe0049M21r, genomic survey
            sequence.
ACCESSION  AQ914559
VERSION    AQ914559.1  GI:6511075
KEYWORDS   GSS.
SOURCE     Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
            1 (bases 1 to 814)
            Wing, R. A. and Dean, R. A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   Location/Qualifiers
            source             1..814
                                /organism="Oryza sativa (japonica cultivar-group)"
                                /mol_type="genomic DNA"
                                /cultivar="Nipponbare"
                                /db_xref="taxon:39947"
                                /clone="nbe0049M21r"
                                /tissue_type="Leaf"
                                /lab_host="E. coli DH10B"
                                /clone_lib="CUGI Rice BAC Library (EcoRI)"
                                /note="Vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
                                Rice is the most important food crop in the world. Half of
                                the world population, especially those inhabiting highly
                                populated areas of the humid tropics and subtropics, rely
                                on rice as their primary source of carbohydrate.
                                Monocotyledonous rice is a diploid plant (2n=24) with a
                                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, the Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center ([www.genome.clemson.edu](http://www.genome.clemson.edu)).

## ORIGIN

Query Match 53.2%; Score 53.2; DB 8; Length 814;  
 Best Local Similarity 78.0%; Pred. No. 6.7e-08;  
 Matches 64; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 7 TGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATGCCGAT 66  
 |||||  
 Db 279 TGGGGGATTTACACCGCATATGGTCACCTCTCAGTACATCTGCTCTGATGCCGAT 338  
 |||||

QY 67 AGTTAAGCCAGTATCTGCTCCC 88  
 |||||

Db 339 AGTTAAGCCAGCCCGCACCC 360  
 |||||

## RESULT 15

BZ051815  
 LOCUS jnr57d03.b1 B.oleracea001 675 bp DNA linear GSS 09-OCT-2002  
 DEFINITION jnr57d03.b1 B.oleracea001 Brassica oleracea genomic, genomic survey sequence.

ACCESSION BZ051815  
 VERSION BZ051815.1 GI:23652690  
 KEYWORDS GSS.

SOURCE  
 ORGANISM Brassica oleracea  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica

REFERENCE  
 AUTHORS Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,  
 Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
 TITLE Whole genome shotgun reads from Brassica oleracea  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 Plate: jnr57 row: d column: 03  
 Seq primer: -21UPPOT forward

Class: shotgun  
 High quality sequence start: 29  
 High quality sequence stop: 94.

## FEATURES

source  
 1..675  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOTw13; Whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea T01000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin. Genomic  
 DNA was provided by Pablo Rabinowicz (CSHL) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."

## ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;  
 Best Local Similarity 75.6%; Pred. No. 7.6e-08;

Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CCGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATGCC 62  
 |||||  
 Db 53 CCGACGATAGTCCCTGGACTAGTTATGGTGCACCTCTCAGTACATCTGCTCTGATGCC 112  
 |||||

QY 63 GCATAGTTAAGCCAGTATCTGCTCCC 88  
 |||||

Db 113 GCATAGTTAAGCCAGCCCGCACCC 138  
 |||||

Search completed: July 14, 2005, 23:22:24  
 Job time : 957.146 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_3674\_3774  
Perfect score: 101  
Sequence: 1 cggctattctttgattta.....acgggaattaattctgtgga 101

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_phi.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_to.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	3986	12	PCDNA3ZEO
2	101	100.0	4597	6	AX060344 Sequence
3	101	100.0	5070	6	AX234391 Sequence
4	101	100.0	5082	6	AX1754 Sequence 10
5	101	100.0	5082	6	BD085110
6	101	100.0	5432	6	BD234590
7	101	100.0	5432	6	AX026821 Sequence
8	101	100.0	5446	6	BD195386
9	101	100.0	5446	6	AX319694
10	101	100.0	5639	12	AX437643
11	101	100.0	5651	6	AX211282
12	101	100.0	5651	6	AX349366
13	101	100.0	5731	6	AX202478
14	101	100.0	5995	6	AX685746
15	101	100.0	6084	12	GA575208
16	101	100.0	6109	12	TRU90717
17	101	100.0	6148	6	BD181637
18	101	100.0	6148	6	AX342685
19	101	100.0	6149	6	BD181638

20	101	100.0	6149	6	AX342686	AX342686 Sequence
21	101	100.0	6180	6	AX207724	AX207724 Sequence
22	101	100.0	6186	6	AX211281	AX211281 Sequence
23	101	100.0	6186	6	AX349365	AX349365 Sequence
24	101	100.0	6195	6	BD168975	BD168975 Method of
25	101	100.0	6213	6	AX211283	AX211283 Sequence
26	101	100.0	6213	6	AX349369	AX349369 Sequence
27	101	100.0	6232	6	AR409004	AR409004 Sequence
28	101	100.0	6238	6	BD168966	BD168966 Method of
29	101	100.0	6253	6	AR031374	AR031374 Sequence
30	101	100.0	6253	6	BD009742	BD009742 Compositi
31	101	100.0	6277	12	AX437644	AX437644 Expressio
32	101	100.0	6331	12	EVPCMVPA1	X96612 Expression
33	101	100.0	6333	12	EVPCMVPA3	X96611 Expression
34	101	100.0	6335	12	EVPCMVPA2	X96610 Expression
35	101	100.0	6338	6	BD134374	BD134374 Peptide 1
36	101	100.0	6338	6	AR428934	AR428934 Sequence
37	101	100.0	6340	6	AX207733	AX207733 Sequence
38	101	100.0	6365	6	AX513181	AX513181 Sequence
39	101	100.0	6394	12	AF416990	AF416990 Synthetic
40	101	100.0	6404	6	BD267665	BD267665 Delivery
41	101	100.0	6411	6	AX207725	AX207725 Sequence
42	101	100.0	6411	6	AX207729	AX207729 Sequence
43	101	100.0	6420	6	BD267666	BD267666 Delivery
44	101	100.0	6436	6	AX207740	AX207740 Sequence
45	101	100.0	6439	6	AR240214	AR240214 Sequence

ALIGNMENTS

RESULT 1  
PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995  
LOCUS Cloning vector pcdna3zео DNA.  
DEFINITION X90639  
ACCESSION X90639.1 GI:949972  
VERSION cloning vector; expression vector; multiple cloning site; Plasmid.  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Peters, H., Hundhausen, T., Kroenke, M. and Marget, M.  
TITLE A new small sized high-level eukaryotic expression vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3986)  
AUTHORS Peters, H.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-1995): H. Peters, Inst. f. Immunologie,  
Michaelisstr. 5, D- 24105 Kiel, FRG  
COMMENT Related sequences: M21295 and K03104.  
FEATURES  
source  
Location/Qualifiers  
1..3986  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
/plasmid="pcDNA3ZEO"  
1..2125  
misc\_feature  
/note="cloning vector (pcDNA3). (Invitrogen)"  
889..994  
misc\_feature  
/note="multiple cloning site (MCS)"  
2126..2796  
misc\_feature  
/note="cloning vector (PZeoSV) (Invitrogen)"  
2797..3986  
misc\_feature  
/note="cloning vector (pcDNA3)"

ORIGIN

Query Match 100.0%; Score 101; DB 12; Length 3986;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTCGCCCTATTGGTTAAAAAATG 60  
|||||

Db 1651 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
|||||  
Db 1711 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1751

RESULT 2  
AX060344  
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 3 from Patent WO0078358.  
ACCESSION AX060344  
VERSION AX060344.1 GI:12405832  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Chen, W.  
TITLE Hyaluronic acid microspheres for sustained gene transfer  
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;  
The Collaborative Group, Ltd. (US)  
FEATURES  
source  
1. 4597  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 4597;  
Best Local Similarity 100.0%; Pred. No. 4.9e-15;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60  
|||||

Db 2227 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 2286  
|||||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
|||||

Db 2287 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 2327  
|||||

RESULT 3  
AX234391  
LOCUS AX234391 5070 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 41 from Patent WO0162942.  
ACCESSION AX234391  
VERSION AX234391.1 GI:15593392  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Altalo, K.M. and Jeltsch, M.M.  
TITLE Materials and methods involving hybrid vascular endothelial growth factor dnas and proteins and screening methods for modulators  
JOURNAL Patent: WO 0162942-A 41 30-AUG-2001;  
LUDWIG INSTITUTE FOR CANCER RESEARCH (US); Licentia OY (FI)  
FEATURES  
source  
1. 5070  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pSecTag1 Vector"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5070;  
Best Local Similarity 100.0%; Pred. No. 4.8e-15;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60  
|||||

Db 1693 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 1752

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
|||||

Db 1753 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1793  
|||||

RESULT 4  
A91754  
LOCUS A91754 5082 bp DNA circular PAT 22-JAN-2000  
DEFINITION Sequence 10 from Patent WO9824810.  
ACCESSION A91754  
VERSION A91754.1 GI:6740671  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 5082)  
AUTHORS Bogaert, T.A. and Deraeymaeker, M.  
TITLE VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS  
JOURNAL Patent: WO 9824810-A 10 11-JUN-1998;  
BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEYMAEKER MARC (BE)  
FEATURES  
Location/Qualifiers  
1. 5082  
source  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 100.0%; Score 101; DB 6; Length 5082;  
Best Local Similarity 100.0%; Pred. No. 4.8e-15;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60  
|||||

Db 3321 CGGTCATCTTTGATTATAAGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 3380  
|||||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
|||||

Db 3381 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 3421  
|||||

RESULT 5  
BD085110  
LOCUS BD085110 5082 bp DNA linear PAT 27-AUG-2002  
DEFINITION Vertebrate homologues of UNC-53 protein of C elegans.  
ACCESSION BD085110  
VERSION BD085110.1 GI:22630720  
KEYWORDS JP 2001522222-A/8.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 5082)  
AUTHORS Platteeuw, C.J., Arjol, C.M.B., Deraeymaeker, M., Verhasselt, P., Pujol, N.J.R., Luc, Maertens, J.S., Luyten, W., Geerts, H., Vandekerckhove, J.S., Geysen, J. and Bogaert, T.A.O.E.  
TITLE Vertebrate homologues of UNC-53 protein of C elegans  
JOURNAL Patent: JP 2001522222-A 8 13-NOV-2001;  
JANSSEN PHARMACEUTICA NV  
COMMENT  
OS Unidentified  
PN JP 2001522222-A/8  
PD 13-NOV-2001  
PF 03-DEC-1997 JP 1998525231  
PR 04-DEC-1996 GB 9625283.8  
PI CHRIST JULES PLATTEEUW, CARLOS MANUEL BUESA ARJOL, MARC PI DERAEYMAEKER,  
PI PETER VERHASSELT, NATHALIE JEANNE RAYMONDE PUJOL, LUC PI JACQUES SIMON MAERTENS,  
PI WALTER LUYTEN, HUGO GEERTS, JOEL STEFAAN VANDEKERCKHOVE, JOHAN PI 'GEYSEN',  
PI THIERRY ANDRE OLIVIER EDDY BOGAERT  
PC C12N15/12, C12N5/10, C12N15/85, C07K14/435, C07K16/18, A61K38/17, A61K49/00,

PC C12Q1/02.G01N33/53  
 CC Strandedness: Double;  
 CC Topology: Circular;  
 CC Vertebrate homologues of UNC-53 protein of C elegans FH Key  
 FT source 1..5082 /organism='Unidentified'.  
 FT Location/Qualifiers

## FEATURES

source  
 1..5082  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5082;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-15;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 3321 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 3380  
 QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 101  
 Db 3381 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 3421

## RESULT 6

LOCUS BD234590 5432 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Screening assay of Abeta-peptide.  
 ACCESSION BD234590  
 VERSION BD234590.1 GI:33044360  
 KEYWORDS JP 2002531141-A/2.  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.

## REFERENCE

1 (bases 1 to 5432)  
 Peraus, G.

## AUTHORS

Screening assay of Abeta-peptide

TITLE Screening assay of Abeta-peptide

JOURNAL Patent: JP 2002531141-A 2 24-SEP-2002;

COMMENT AVENTIS PHARMA DEUTSCHLAND GMBH

OS Artificial Sequence

PN JP 2002531141-A/2

PD 24-SEP-2002

PF 27-NOV-1999 JP 2000586944

PR 07-DEC-1998 DE 198 56 261.6

PI GISELA PERAUS

PC C12N15/09, A01K67/033, A61K45/00, A61P25/28, C12N1/15, C12N1/19, PC

C12N1/21,

PC C12N5/10, C12Q1/37, C12Q1/68, C12N15/00, C12N5/00 CC Description

Of Artificial Sequence: Mutagen

FT Key Location/Qualifiers

FT source 1..5432 /organism='Artificial Sequence'.  
 FT Location/Qualifiers

source 1..5432  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

FEATURES

source

## ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5432;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 1637 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 1696  
 QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 101  
 Db 1697 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 1737

## RESULT 7

LOCUS AX026821 5432 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 9 from Patent DE19856261.  
 ACCESSION AX026821  
 VERSION AX026821.1 GI:10187947  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.

## REFERENCE

1

AUTHORS Peraus, G.

JOURNAL Patent: DE 19856261-C 9 30-MAR-2000;

HOECHST MARION ROUSSEL DE GMBH (DE)

FEATURES Location/Qualifiers

source 1..5432  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Mutagen"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5432;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 1637 CGGTCATCTTTGATTATTAAGGATTTGGGATTTGGCCCTATTGGTTAAAAAATG 1696  
 QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 101  
 Db 1697 AGCTGATTTACAAAAATTTAACGCGAATTAATCTCTGTGGA 1737

## RESULT 8

LOCUS BD195386 5446 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Composition and methods for administering Pneumococcal DNA.  
 ACCESSION BD195386  
 VERSION BD195386.1 GI:33005156  
 KEYWORDS JP 2002514061-A/3.  
 SOURCE unidentified  
 ORGANISM unclassified.

## REFERENCE

1 (bases 1 to 5446)

Briles, D.E., McDaniel, L.S. and Curriel, D.T.

Composition and methods for administering Pneumococcal DNA

TITLE Patent: JP 2002514061-A 3 14-MAY-2002;

JOURNAL UNIVERSITY OF ALABAMA AT BIRMINGHAM

COMMENT OS Unidentified

PN JP 2002514061-A/3

PD 14-MAY-2002

PF 04-DEC-1997 JP 1998525895

PI 04-DEC-1996 US 08/759505

CC Strandedness: Single;

CC Topology: Linear;

CC Composition and methods for administering Pneumococcal DNA FH

FT Key Location/Qualifiers

FT source 1..5446 /organism='Unidentified'.  
 FT Location/Qualifiers

source 1..5446  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5446;

```
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1651 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1711 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 1751

RESULT 9
AX319694 AX319694 5446 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 5 from Patent WO0181614.
DEFINITION AX319694
ACCESSION AX319694
VERSION AX319694.1 GI:17901350
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Leng,J.
TITLE Cell proliferation assay
JOURNAL Patent: WO 0181614-A 5 01-NOV-2001;
Chemicon International (US)
FEATURES
source
1. 5446
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pcDNA3 vector sequence"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1651 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1711 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 1751

RESULT 10
AY437643 AY437643 5639 bp DNA circular SYN 10-NOV-2003
LOCUS Expression vector pcGlobin 2, complete sequence.
DEFINITION AY437643
ACCESSION AY437643
VERSION AY437643.1 GI:38155839
KEYWORDS Expression vector pcGlobin 2
SOURCE Expression vector pcGlobin 2
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS Ro.H., Kim,E.J. and Rhee,M.
TITLE A new vector system, pcGlobin 2 for in vitro synthesized RNA
injection into zebrafish embryos
JOURNAL Unpublished
REFERENCE 2
AUTHORS Ro.H., Kim,E.J. and Rhee,M.
TITLE Direct Submision
JOURNAL Submitted (14-OCT-2003) Department of Biology, College of Natural
Sciences, Chungnam National University, 305-764, Daejeon 305-764,
Korea
FEATURES
source
1. 5639
/organism="Expression vector pcGlobin 2"
/mol_type="Other DNA"
```

```
/db_xref="taxon:254096"
/note="eukaryotic expression vector for zebrafish embryo
microinjection; derivative of pcDNA3"
complement(4643..5503)
/codon_start=1
/product="beta-lactamase"
/protein_id="AAR12689.1"
/db_xref="GI:38155840"
/translations="MSIQHFRVALIIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IEIDLNSGKILESFRPEERPPMWSFKVLCCGAVLSRIDAGQEQLRRIRHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFIHNNGDHTRL
DRWEPELNEAIPNDERDTTPVAMATLRKLLTGELLTLASRQQLIDWMEADKVGPL
LRSLPAGWFIADKSGAGERSGIILAAALGPDGKPSRIVVIYITGSGQATWDERNRQIA
EIGASLIKHW"

ORIGIN

Query Match 100.0%; Score 101; DB 12; Length 5639;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1844 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1903

Qy 61 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1904 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 1944

RESULT 11
AX211282 AX211282 5651 bp DNA linear PAT 06-SEP-2001
LOCUS Sequence 6 from Patent WO0158493.
DEFINITION AX211282
ACCESSION AX211282
VERSION AX211282.1 GI:15523691
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Schambye,H.T., Andersen,K.V., van den Hazel,B., Christiansen,J. and
Jeppesen,C.B.
TITLE Conjugates of follicle stimulating hormones
JOURNAL Patent: WO 0158493-A 6 16-AUG-2001;
Maxygen Aps (DK)
FEATURES
source
1. 5651
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Coding sequence for human FSH-beta"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 2273 CGGTCATCTTTTGATTATTAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 2332

Qy 61 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 101
Db 2333 AGCTGATTTACAAAATTTAACGCGAATTAATCTGTGGA 2373

RESULT 12
AX349366 AX349366 5651 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 4 from Patent WO0202597.
DEFINITION AX349366
ACCESSION AX349366
VERSION AX349366.1 GI:18615329
```



```
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .5651
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1231. .1617
/note="Coding sequence for human FSH-beta"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 60
DB 2273 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 2332
QY 61 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCCTGTGGA 101
DB 2333 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCCTGTGGA 2373
RESULT 13
LOCUS
AX202478
DEFINITION
Sequence 66 from Patent WO0152620.
ACCESSION
AX202478
VERSION
AX202478.1 GI:15392206
KEYWORDS
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS
Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE
Methods and compositions to modulate expression in plants
JOURNAL
Patent: WO 0152620-A 66 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source
Location/Qualifiers
1. .5731
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="2C7-SID"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5731;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 60
DB 2353 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 2412
QY 61 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCCTGTGGA 101
DB 2413 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCCTGTGGA 2453
RESULT 14
LOCUS
AX685746
DEFINITION
Sequence 5 from Patent WO02102854.
ACCESSION
AX685746
```

```
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .5995
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
<963. .1670
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD86574.1"
/db_xref="GI:29371752"
/translation="DPEPKSCDKHTPCPCPAPPELLCGSPVFLPPKPKDMLISRT
PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTIRRVSVLTFLHQDWL
NGKEYCKCKVSNKALPAPIEKTIYAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGF
YPSDIAVWESNGQPENNYKTTPPVLDSGSEFFLYSKLTVDKSRWQQGNVSPFCSVMHE
ALHNHYTKSLSLSPGK"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5995;
Best Local Similarity 100.0%; Pred. No. 4.6e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 60
DB 2497 CGGTCCTATCTTTGATTATTAAGGATTTTGGGATTTTCGCCCTATTGCTTAAAAAATG 2556
QY 61 AGCTGATTTTACAAAAATTTAACCGCAATTAATTCCTGTGGA 101
DB 2557 AGCTGATTTTACAAAAATTTAACCGCAATTAATTCCTGTGGA 2597
RESULT 15
LOCUS
GGA575208
DEFINITION
Expression vector pCLGPPA.
ACCESSION
AJ575208
VERSION
AJ575208.1 GI:32451228
KEYWORDS
Expression vector pCLGPPA
Expression vector pCLGPPA
other sequences; artificial sequences; vectors.
ORGANISM
REFERENCE
1
AUTHORS
Scaal,M., Gros,J., Lesbros,C. and Marcelle,C.
TITLE
In ovo electroporation of avian somites
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 6084)
AUTHORS
Marcelle,C.
TITLE
Direct Submision
JOURNAL
Submitted (13-JUN-2003) Marcelle C., Lgpd, Institut de Biologie du
Developpement, Campus de Luminy Case 907 F-13288 Marseille, 13288,
FRANCE
FEATURES
source
Location/Qualifiers
1. .6084
/organism="Expression vector pCLGPPA"
/mol_type="other DNA"
/db_xref="taxon:236984"
/focus
12. .397
/organism="unidentified cytomegalovirus"
/mol_type="other DNA"
/db_xref="taxon:205912"
398. .1652
source
```

```

/organism="Gallus gallus"
/mol_type="other DNA"
/db_xref="taxon:9031"
1653..1733
/organism="Oryctolagus cuniculus"
/mol_type="other DNA"
/db_xref="taxon:9986"
1782..1996
/organism="Bos taurus"
/mol_type="other DNA"
/db_xref="taxon:9913"
2442..2836
/organism="Simian virus 40"
/mol_type="other DNA"
/db_xref="taxon:10633"
2908..3723
/organism="Aequorea victoria"
/mol_type="other DNA"
/db_xref="taxon:6100"
12..399
740..887
/gene="beta actin"
740..887
/gene="beta actin"
1734..1776
/note="multiple cloning site; MCS
under control of CMV enhancer/chicken beta-actin promoter"
1782..1996
/note="bovine growth hormone"
2442..2836
2908..3723
/note="eGFP
under control of SV40 promoter"
3757..3887
/note="late"

```

## ORIGIN

```

Query Match      100.0%; Score 101; DB 12; Length 6084;
Best Local Similarity 100.0%; Pred. No. 4.6e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGGTCATTCTTTTGATTATTAAGGGATTTCGGGATTTCGGCCTATTGGTTAAAAAATG 60
      |||
Db      2398  CGGTCATTCTTTTGATTATTAAGGGATTTCGGGATTTCGGCCTATTGGTTAAAAAATG 2457

QY      61  AGCTGATTTAACAAAAATTAAACGCGAATTAAATCTGTGGA 101
      |||
Db      2458  AGCTGATTTAACAAAAATTAAACGCGAATTAAATCTGTGGA 2498

```

Search completed: July 14, 2005, 14:03:21  
Job time : 757.618 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds

(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_3674\_3774

Perfect score: 101

Sequence: 1 cggctattcttttgattta.....acoggaattaattctgtgga 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	3482	2	ADH11353
2	101	100.0	4597	4	AAT24901
3	101	100.0	4639	6	AAD339652
4	101	100.0	5015	10	ADB33528
5	101	100.0	5070	4	AAS12839
6	101	100.0	5082	2	ADH11417
7	101	100.0	5173	6	ABK88869
8	101	100.0	5173	12	ADE83792
9	101	100.0	5173	12	ADO06721
10	101	100.0	5218	12	ADM97811
11	101	100.0	5302	12	ADI34681
12	101	100.0	5425	2	ADH11233
13	101	100.0	5431	6	ABN86685
14	101	100.0	5431	10	ADE21866
15	101	100.0	5431	12	ADO05277
16	101	100.0	5432	3	AAZ289476
17	101	100.0	5446	2	AAV38297
18	101	100.0	5446	6	AAI18619
19	101	100.0	5446	6	ABL53540
20	101	100.0	5446	12	ADN36314

21	101	100.0	5458	6	ABL58494
22	101	100.0	5458	6	ABL58493
23	101	100.0	5543	6	ABK88868
24	101	100.0	5543	12	ADE83791
25	101	100.0	5543	12	ADO06720
26	101	100.0	5614	6	ABL58489
27	101	100.0	5614	6	ABL58490
28	101	100.0	5651	5	AAI166195
29	101	100.0	5651	6	ABK40237
30	101	100.0	5695	6	ABL58492
31	101	100.0	5695	6	ABL58491
32	101	100.0	5695	8	ABT40262
33	101	100.0	5695	8	ADA89054
34	101	100.0	5695	10	ADG74306
35	101	100.0	5731	4	AAI11615
36	101	100.0	5864	6	AAI44423
37	101	100.0	5864	6	AAI44424
38	101	100.0	6082	8	AAI56212
39	101	100.0	6082	8	AAI56211
40	101	100.0	6082	8	AAI56210
41	101	100.0	6085	8	AAI56213
42	101	100.0	6094	8	AAI56215
43	101	100.0	6097	8	AAI56214
44	101	100.0	6100	6	ABK96469
45	101	100.0	6135	6	ABK96470

## ALIGNMENTS

### RESULT 1

ADH11353

ID ADH11353 standard; DNA; 3482 BP.

XX

AC ADH11353;

XX

DT 11-MAR-2004 (first entry)

XX

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX

KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;

KW cell shape regulator; cell motility regulator; cell migration;

KW cell behaviour regulator; phenotype; signal transduction pathway;

KW signal transducing protein; phenotypic; signal integrator protein;

KW neuronal regeneration; revascularisation; wound healing;

KW chronic neurodegenerative disease; acute traumatic injury;

KW fibrotic disease; gene; ds.

XX

OS Unidentified.

XX

PN WO9824810-A2.

XX

PD 11-JUN-1998.

XX

PF 03-DEC-1997; 97WO-EP006956.

XX

PR 04-DEC-1996; 96GB-00025283.

XX

PA (JANC ) JANSSEN PHARM NV.

XX

PI Platteauw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

PI Geyse J, Bogaert TAOE,

XX

WPI; 1998-362411/31.

DR P-PSDB; ADH11354.

XX

PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

PT promoting neuronal regeneration, treating chronic neuro-degenerative

PT diseases or acute traumatic injuries.

XX

PS Disclosure; Page 414-417; 479pp; English.

XX

CC The present invention describes a vertebrate protein homologue of an UNC-53 protein of *Caenorhabditis elegans* or a functional equivalent, CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a nucleic acid which hybridises to the cDNA of (1); (3) vector comprising the cDNA as in (1); (4) a host cell containing the vector as in (3); (5) a transgenic cell, tissue or animal comprising the vector as in (3); (6) a compound identified as an enhancer or inhibitor of the regulation of cell shape, motility, or the direction of cell migration for use as a therapeutic; (7) a method for determination of whether a protein is an inhibitor or enhancer of regulation of cell behaviour, growth, shape or motility or the direction of migration by contacting a host cell CC expressing a homologue of UNC-53 and determining a change of phenotype; CC (8) a method for identification of vertebrate homologues of *C. elegans* unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to a DNA library; and (9) a method for identification of a protein which is active in the signal transduction pathway of a cell of which a vertebrate CC homologue of UNC-53 is a component comprising: (i) contacting an extract of a cell with an antibody to the UNC-53 homologue; (ii) identifying an CC antibody/homologue complex; and (iii) analysing such a complex to CC identify any non-antibody protein bound to the complex. UNC-53 is a CC signal transducing or signal integrator protein involved in controlling CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate CC homologues of UNC-53 can be used to promote neuronal regeneration, CC revascularisation or wound healing, to treat chronic neurodegenerative CC diseases or acute traumatic injuries or fibrotic diseases. The present CC sequence is used in the exemplification of the present invention.

XX Sequence 3482 BP; 767 A; 956 C; 913 G; 846 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 3482;

Best Local Similarity 100.0%; Pred. No. 4.1e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60  
|||||  
Db 1721 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 1780

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101  
|||||

Db 1781 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1821

RESULT 2

AAF24901

ID AAF24901 standard; DNA; 4597 BP.

XX AAF24901;

DT 20-APR-2001 (first entry)

XX Nucleotide sequence of the plasmid pCDNA3.1/GS.

DE Microsphere; dihydrazide; hyaluronic acid; inflammatory response;  
KW myocardial ischemia; cardiac angiogenesis; haemophilia;  
KW vascular endothelial growth factor; VEGF; ss.

OS Synthetic.

XX WO200078358-A2.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US016837.

XX 18-JUN-1999; 99US-0140260P.

XX (COLL-) COLLABORATIVE GROUP LTD.

XX Chen W;

XX WPI; 2001-071363/08.

XX

PT Hyaluronic acid micro spheres for use in gene therapy of myocardial  
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic  
PT acids crosslinked to nucleic acids.

XX Example 1; Page 36-38; 38pp; English.

XX The specification describes a microsphere comprising dihydrazide  
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The  
CC microspheres cause reduced inflammatory responses, and have increased  
CC safety and biodegradability. The microspheres are useful for transfecting  
CC a cell of a subject and for treating a subject having myocardial  
CC ischemia, by increasing cardiac angiogenesis. They are also useful for  
CC treating haemophilia. The present sequence represents the plasmid  
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is  
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a  
CC vascular endothelial growth factor (VEGF)

XX Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 4597;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60  
|||||

Db 2227 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 2286

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101  
|||||

Db 2287 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 2327

RESULT 3

AAD39652

ID AAD39652 standard; DNA; 4639 BP.

XX AAD39652;

XX 22-OCT-2002 (first entry)

XX Human small nuclear RNA (snRNA) DNA.

XX Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;  
KW transgenic animal; ds.

XX Homo sapiens.

XX US2002058287-A1.

XX 16-MAY-2002.

XX 12-MAR-2001; 2001US-00804481.

XX 10-MAR-2000; 2000US-0188304P.

XX (WHEED ) WHITEHEAD INST BIOMEDICAL RES.

XX Graaf DD, Lander ES;

XX WPI; 2002-499510/53.

XX New recombinant vector containing sequence for small nuclear RNA, useful  
PT e.g. for identifying variant snRNA that suppresses expression of  
PT transcription products.

XX Disclosure; Fig 1; 18pp; English.

XX The invention relates to a recombinant vector which comprises DNA,  
CC consisting of an insertion cassette contained between at least two  
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is  
CC used to identify snRNA modifications that inhibit expression of  
CC transcription products (and the identified snRNA are used to suppress  
CC expression) for delivering antisense sequences to the nucleus and to

CC create transgenic animals. The present DNA sequence is human snRNA, UI  
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 6; Length 4639;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGGCCCTATTGGTTAAAAAATG 60  
Db 1261 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGGCCCTATTGGTTAAAAAATG 1320  
QY 61 AGCTGATTTTAAACAAAATTTACGCGAATTAATCTGTGGA 101  
Db 1321 AGCTGATTTTAAACAAAATTTACGCGAATTAATCTGTGGA 1361  
RESULT 4  
ADB33528  
ID ADB33528 standard; DNA; 5015 BP.  
XX AC ADB33528;  
XX DT 04-DEC-2003 (first entry)  
XX DE Expression vector nucleotide sequence SEQ ID NO:27.  
XX KW fusion protein; amyloid precursor protein; APP; transcription factor;  
KW neurotrophic; neuroprotective; APP inhibitor;  
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;  
KW gamma-secretase; human; gene; ds.  
XX OS Synthetic.  
XX OS Homo sapiens.  
XX PN WO2003072041-A2.  
XX PD 04-SEP-2003.  
XX PF 23-FEB-2003; 2003WO-US005458.  
XX PR 27-FEB-2002; 2002US-0360274P.  
XX PA (MERI ) MERCK & CO INC.  
XX PI Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;  
PI Miller MD, Register B, Shi X, Simon AJ, Zuck PD;  
XX WPI; 2003-689968/65.  
XX DNA encoding a fusion protein of amyloid precursor protein, useful in  
PT screening for anti-Alzheimer agents, comprises a fused transcription  
PT factor.  
XX PS Disclosure; Fig 32B-F; 193pp; English.  
XX The present invention describes a DNA molecule (I) that encodes a fusion  
CC protein (FP) comprising: (i) an amino acid sequence of amyloid precursor  
CC protein (APP), either the wild type, Swedish or NREV versions; and (ii) a  
CC transcription factor (TF), fused in frame to the C-terminus of (i). Also  
CC described: (1) an expression vector containing (I); (2) a eukaryotic cell  
CC containing (I); and (3) methods for identifying a compound (A) that  
CC inhibits processing of APP, using the cells of (2). (1) has neurotropic and  
CC neuroprotective activities. (I) can be used to produce eukaryotic cells  
CC that express FP and are useful in screening for agents that inhibit  
CC processing of APP. The agents are potentially useful for the treatment or  
CC prevention of Alzheimer's disease. Cells that express FP can screen for  
CC inhibitors of: (a) beta- and gamma-secretases; and (b)  
CC cytoplasmic/extracellular APP signaling in a single assay. Cell-based  
CC assays may be free of interference from alpha-secretase activity and are  
CC homogeneous (no chromatography, immunoprecipitation or washing required)  
CC so well suited to high-throughput screening. The present sequence  
CC represents a plasmid nucleotide sequence from the present invention.

XX SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 10; Length 5015;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGGCCCTATTGGTTAAAAAATG 60  
Db 1637 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGGCCCTATTGGTTAAAAAATG 1696  
QY 61 AGCTGATTTTAAACAAAATTTACGCGAATTAATCTGTGGA 101  
Db 1697 AGCTGATTTTAAACAAAATTTACGCGAATTAATCTGTGGA 1737  
RESULT 5  
AAS12839  
ID AAS12839 standard; DNA; 5070 BP.  
XX AC AAS12839;  
XX DT 21-NOV-2001 (first entry)  
XX DE DNA sequence of pSecTagI vector used to express VEGF-A/VEGF-C hybrids.  
XX KW Human; vascular endothelial growth factor; VEGF-A; vasculogenesis;  
KW angiogenesis; blood vessel; cancer; proliferative retinopathy; psoriasis;  
KW age-related macular degeneration; rheumatoid arthritis; cardiovascular;  
KW pSecTagI; cyclic; circular; ds.  
XX OS Synthetic.  
XX PN WO200162942-A2.  
XX PD 30-AUG-2001.  
XX PF 26-FEB-2001; 2001WO-US006113.  
XX PR 25-FEB-2000; 2000US-0185205P.  
XX PR 18-MAY-2000; 2000US-0205331P.  
XX PA (LUDW-) LUDWIG INST CANCER RES.  
XX PA (LICN ) LICENTIA OY.  
XX PI Alitalo K, Jeltsch MM;  
XX WPI; 2001-536640/59.  
XX PT Polypeptides that bind cellular receptors for vascular endothelial growth  
PT factors, polynucleotides encoding them.  
XX PS Example 1; Page 173-176; 261pp; English.  
XX The present invention relates to polypeptides that bind cellular  
CC receptors for vascular endothelial growth factors (VEGFs), the  
CC polynucleotides encoding them, and their use for identifying agents that  
CC modulate interactions between VEGFs and their receptors. VEGFs and their  
CC receptors play an important role in vasculogenesis, the development of  
CC the embryonic vasculature from early differentiating endothelial cells  
CC and angiogenesis, the process of forming new blood vessels from pre-  
CC existing ones. Modulators of interactions between VEGF and its receptors  
CC may be used to treat dysfunction of the endothelial cell regulatory  
CC system. Such disorders include cancers, abnormal angiogenesis,  
CC proliferative retinopathies, age-related macular degeneration, rheumatoid  
CC arthritis and psoriasis. The polypeptides of the invention exhibit unique  
CC receptor binding profiles compared to known naturally occurring VEGFs.  
CC The present DNA sequence for pSecTagI vector is used to express VEGF-  
CC A/VEGF-C hybrids in the methods of the present invention  
XX SQ Sequence 5070 BP; 1186 A; 1308 C; 1288 G; 1288 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 4; Length 5070;

```

Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGGTTAAAAAATG 60
Db 1693 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGGTTAAAAAATG 1752

QY 61 AGCTGATTTACAAAAATTTACGCGAATTAATTCGTGGA 101
Db 1753 AGCTGATTTACAAAAATTTACGCGAATTAATTCGTGGA 1793

RESULT 6
ADH11417
ID ADH11417 standard; DNA; 5082 BP.
AC ADH11417;
DT 11-MAR-2004 (first entry)
XX
DE Plasmid pCB201 nucleotide sequence.
XX
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; human; gene; ds.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1028..2149
FT /tag= a
XX
PN WO9824810-A2.
XX
PD 11-JUN-1998.
XX
PF 03-DEC-1997; 97WO-EP006956.
XX
PR 04-DEC-1996; 96GB-00025283.
XX
PA (JANC ) JANSSEN PHARM NV.
XX
PI Plattreuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
PI Pujol NTR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geysen J, Bogaert TAOE;
XX
DR WPI; 1998-362411/31.
DR P-PSDB; ADH11424.
XX
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
PT promoting neuronal regeneration, treating chronic neuro-degenerative
PT diseases or acute traumatic injuries.
XX
PS Claim 96; SEQ ID NO 10; 479pp; English.
XX
CC The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of Caenorhabditis elegans or a functional equivalent,
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;

```

(8) a method for identification of vertebrate homologues of C. elegans unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to a DNA library; and (9) a method for identification of a protein which is active in the signal transduction pathway of a cell of which a vertebrate homologue of UNC-53 is a component comprising: (i) contacting an extract of a cell with an antibody to the UNC-53 homologue; (ii) identifying an antibody/homologue complex; and (iii) analysing such a complex to identify any non-antibody protein bound to the complex. UNC-53 is a signal transducing or signal integrator protein involved in controlling directionality of cell migration and cell shape in C. elegans. Vertebrate homologues of UNC-53 can be used to promote neuronal regeneration, revascularisation or wound healing, to treat chronic neurodegenerative diseases or acute traumatic injuries or fibrotic diseases. The present sequence is used in the exemplification of the present invention.

Sequence 5082 BP; 1164 A; 1365 C; 1311 G; 1242 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5082;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGGTTAAAAAATG 60  
Db 3321 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGGTTAAAAAATG 3380

QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATTCGTGGA 101  
Db 3381 AGCTGATTTACAAAAATTTAACGCGAATTAATTCGTGGA 3421

RESULT 7  
ABK88869  
ID ABK88869 standard; DNA; 5173 BP.  
AC ABK88869;  
XX  
DT 07-AUG-2003 (revised)  
DT 07-OCT-2002 (first entry)  
XX  
DE Topoisomerase vector pcDNA6.2/V5/GWD-TOPO.  
XX  
KW ds; topoisomerase recognition site; topoisomerase; pcDNA6.2/V5/GWD-TOPO;  
KW PENTR-DT(sc); pcDNA-DEST41; PENTR/D-TOPO; PENTR/SD/D-TOPO;  
KW pcDNA3.2/V5/GWD-TOPO; pcDNA6.2/V5/GWD-TOPO; recombinational cloning;  
KW gene targeting; mutation; cyclic; circular; vector.  
XX  
OS Escherichia coli.  
OS Viruses.  
OS Human cytomegalovirus.  
OS Synthetic.  
XX  
PN WO200246372-A1.  
XX  
PD 13-JUN-2002.  
XX  
PF 07-DEC-2001; 2001WO-US045773.  
XX  
PR 08-DEC-2000; 2000US-0254510P.  
PR 11-DEC-2000; 2000US-00732914.  
PR 14-SEP-2001; 2001US-0318902P.  
PR 28-SEP-2001; 2001US-0326092P.  
PR 27-NOV-2001; 2001US-0333124P.  
XX  
PA (INVI-) INVITROGEN CORP.  
XX  
PI Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;  
PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;  
XX  
DR WPI; 2002-519662/55.  
XX  
PT New isolated nucleic acid molecule comprises one or more recombination  
PT sites and one or more topoisomerase recognition sites and/or one or more  
PT topoisomerases, useful in recombinational cloning.

XX Disclosure; Fig 25B-25C; 324pp; English.

XX The invention relates to an isolated nucleic acid molecule (1)

XX comprising: (a) one or more recombination sites; and (b) one or more

XX topoisomerase recognition sites and/or one or more topoisomerases. Also

XX included are a vector comprising the nucleic acid, a vector chosen from

XX pCDNAGWDT(sc), pENTR-DT(sc), pCDNA-DEST41, pENTR/D-TOPO, pENTR/SD/D-TOPO,

XX pCDNA3, 2/V5/GWD-TOPO or pCDNA6.2/V5/GWD-TOPO, a host cell comprising the

XX nucleic acid or vectors and an in vitro method of cloning a nucleic acid

XX molecule involving: (a) obtaining a first nucleic acid molecule to be

XX cloned; (b) mixing the first nucleic acid molecule to be cloned in vitro

XX with a second nucleic acid molecule comprising at least a first

XX topoisomerase recognition site flanked by at least a first recombination

XX site, and at least a second topoisomerase recognition site flanked by at

XX least a second recombination site, where the first and second

XX recombination sites do not recombine with each other, and at least one

XX topoisomerase; and (c) incubating the mixture under conditions such that

XX the first nucleic acid molecule to be cloned is inserted into the second

XX nucleic acid molecule between the first and second topoisomerase

XX recognition sites, thereby producing a first product molecule comprising

XX the first nucleic acid molecule to be cloned between the first and second

XX recombination sites. The method is useful for cloning a nucleic acid

XX molecule. The nucleic acid (1) is useful in methods for recombinational

XX cloning and facilitates construction of gene targeting nucleic acid

XX molecules or vectors which may be used to knockout or mutate a sequence

XX or gene of interest, particularly genes or sequences within a host or

XX host cells such as animal, plant, etc. Thus the nucleic acid is most

XX preferably used for targeting or mutating a sequence of gene. The present

XX sequence is the topoisomerase site-containing vector pCDNA6.2/V5/GWD-

XX TOPO. (Updated on 07-AUG-2003 to correct OS field.)

SQ Sequence 5173 BP; 1247 A; 1338 C; 1274 G; 1305 T; 0 U; 9 Other;

Query Match 100.0%; Score 101; DB 6; Length 5173;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1747 CGGCTATTCTTTGATTATTAAGGATTTGGGATTTTCGGCTATTGTTAAAAAATG 1806

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101

Db 1807 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1847

RESULT 8

ADE83792

ID ADE83792 standard; DNA; 5173 BP.

XX ADE83792;

XX ADE83792;

DT 29-JAN-2004 (first entry)

XX Plasmid pCDNA6.2/V5/GWD-TOPO.

XX recombination site; topoisomerase recognition site; topoisomerase;

XX transfection; two-hybrid assay; ds; plasmid; cyclic; circular;

XX pCDNA6.2/V5/GWD-TOPO.

XX Unidentified.

XX US2003186233-A1.

XX 02-OCT-2003.

XX 07-DEC-2001; 2001US-00005876.

XX 08-DEC-2000; 2000US-0254510P.

XX 11-DEC-2000; 2000US-00732914.

XX 21-MAY-2001; 2001US-0291972P.

XX 14-SEP-2001; 2001US-0318902P.

PR 28-SEP-2001; 2001US-0326092P.

PR 27-NOV-2001; 2001US-0333124P.

XX (INVI-) INVITROGEN CORP.

XX Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;

PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;

XX WPI; 2004-031998/03.

XX Isolated nucleic acid molecule useful as vector in host cell comprises at

XX least one recombination site, and at least one topoisomerase recognition

XX site and/or at least one topoisomerase.

XX Disclosure; Fig 25B-C; 136pp; English.

XX The invention describes an isolated nucleic acid molecule comprising at

XX least one recombination site, and at least one topoisomerase recognition

XX site and/or at least one topoisomerase. The isolated nucleic acid

XX molecule is used as a vector in a host cell. It can be used directly for

XX transfecting a cell, or as a template for performing amplification, e.g.

XX PCR, a recombination reaction, an in vitro transcription reaction, or a

XX coupled transcription/translation reaction. The invention allows several

XX nucleic acid molecules to be covalently linked in a predetermined

XX directional orientation. A functional product can be selected in vitro by

XX performing an amplification reaction using primers specific for the

XX termini of the desired covalently linked recombinant nucleic acid

XX molecule. This sequence represents plasmid pCDNA6.2/V5/GWD-TOPO

XX associated with the isolation and analysis of the novel polynucleotide of

XX the invention.

SQ Sequence 5173 BP; 1247 A; 1338 C; 1274 G; 1305 T; 0 U; 9 Other;

Query Match 100.0%; Score 101; DB 12; Length 5173;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1747 CGGCTATTCTTTGATTATTAAGGATTTGGGATTTTCGGCTATTGTTAAAAAATG 1806

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101

Db 1807 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1847

RESULT 9

ADO06721

ID ADO06721 standard; DNA; 5173 BP.

XX ADO06721;

XX ADO06721;

DT 01-JUL-2004 (first entry)

XX Recombinatorial cloning method plasmid vector - pCDNA6.2/V5/GWD-TOPO.

XX vector; recombination site; topoisomerase recognition site;

XX topoisomerase; in vitro cloning; recombinatorial cloning; plasmid;

XX pCDNA6.2/V5/GWD-TOPO; ds.

XX Unidentified.

XX WO2003103600-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-US018036.

XX 05-JUN-2002; 2002US-0385613P.

XX (INVI-) INVITROGEN CORP.

XX Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;

PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;  
XX WPI; 2004-090674/09.  
XX Novel isolated nucleic acid molecule comprising recombination sites,  
PT topoisomerase recognition sites and/or topoisomerases, useful for in  
PT vitro cloning of nucleic acid.  
XX Claim 12; Fig 25; 369pp; English.  
XX The invention comprises an isolated nucleic acid sequence (e.g. a vector)  
CC that contains one or more recombination sites, one or more topoisomerase  
CC recognition sites, and/or one or more topoisomerases. The nucleic acid  
CC sequence of the invention is useful for in vitro cloning (e.g.  
CC recombinatorial cloning) of a nucleic acid molecule. The present DNA  
CC sequence represents a plasmid vector of the invention.  
XX Sequence 5173 BP; 1247 A; 1337 C; 1274 G; 1305 T; 0 U; 10 Other;  
SQ  
Query Match 100.0%; Score 101; DB 12; Length 5173;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
Db 1747 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1806  
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
Db 1807 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTGTGGA 1847  
RESULT 10  
ID ADM97811 standard; DNA; 5218 BP.  
XX ADM97811;  
XX 01-JUL-2004 (first entry)  
XX & X UAS beta-lactamase vector SEQ ID NO: 64.  
XX enzyme; sensor cell; signal transduction detection system; promoter;  
KW targeting sequence; targeted drug; ds; vector.  
XX Synthetic.  
OS Unidentified.  
XX WO2004031415-A2.  
XX 15-APR-2004.  
XX 05-SEP-2003; 2003WO-US028078.  
XX 05-SEP-2002; 2002US-0408297P.  
XX (VERT-) VERTEX PHARM INC.  
XX Whitney MA, Zeh K, Sanders PS;  
XX WPI; 2004-330208/30.  
XX Developing a sensor cell, useful in determining the activity of a target  
PT gene and in developing therapeutic drugs, comprises providing cells  
PT comprising a signal transduction detection system and introducing DNA  
PT construct into cells.  
XX Example 7; Page 231-234; 234pp; English.  
XX The present invention relates to a method of developing a sensor cell,  
CC for determining the activity of a target gene in the cell, which  
CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing

CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
CC vector used in the exemplification of the invention.  
XX Sequence 5218 BP; 1231 A; 1361 C; 1335 G; 1291 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 101; DB 12; Length 5218;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
Db 4624 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 4683  
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
Db 4684 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTGTGGA 4724  
RESULT 11  
ID ADI34681 standard; DNA; 5302 BP.  
XX ADI34681;  
XX 22-APR-2004 (first entry)  
XX Nucleotide sequence of plasmid pcDNA6/Biotag (TM) -D-TOPO.  
KW Recombinational cloning; recombination; topoisomerase;  
KW fusion protein production; ds.  
OS Synthetic.  
XX WO2004005482-A2.  
XX 15-JAN-2004.  
XX 08-JUL-2003; 2003WO-US021339.  
XX 08-JUL-2002; 2002US-0393756P.  
PR 19-JUL-2002; 2002US-0396627P.  
PR 10-OCT-2002; 2002US-0417172P.  
XX (INVI-) INVITROGEN CORP.  
XX Bennett RP;  
XX WPI; 2004-091356/09.  
XX New isolated nucleic acid molecules having one or more recombination  
PT sites and encoding an amino acid sequence tag, useful for recombinational  
PT and/or topoisomerase-mediated cloning methods for producing fusion  
PT proteins.  
XX Example 1; Fig 11A-B; 135pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC one or more recombination sites, and one or more nucleic acid sequences  
CC which encode an amino acid sequence tag. (I) can also comprise one or  
CC more topoisomerase recognition sites and/or one or more topoisomerases.  
CC The amino acid sequence tag is an amino acid sequence that is capable of  
CC being post-translationally modified, and is an amino acid sequence that  
CC is capable of being post-translationally modified by biotinylation,  
CC attachment of 4-phosphopantetheine, attachment of lipolic acid or  
CC attachment of flavins, and is an amino acid sequence that is capable of  
CC being biotinylated, wherein the amino acid sequence that is capable of  
CC being biotinylated is all or a portion of the Klebsiella pneumoniae  
CC oxalacetate decarboxylase a subunit, all or a portion of the  
CC Propionibacterium shermanii transcarboxylase 1.3S subunit, or all or a



CC portion of the Escherichia coli biotin carboxyl carrier protein component  
 CC of acetyl-CoA carboxylase. The methods and compositions of the present  
 CC invention are useful for identifying, concentrating, purifying and/or  
 CC producing fusion proteins that comprise an amino acid sequence tag. The  
 CC nucleic acid molecules can also be used in recombinational cloning and/or  
 CC topoisomerase-mediated cloning methods to produce polynucleotide  
 CC constructs which encode the fusion proteins. The present sequence  
 CC represents the nucleotide sequence of a plasmid pCDNA6/Biotag(TM)-D-TOPO  
 XX Sequence 5302 BP; 1254 A; 1361 C; 1349 G; 1338 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 101; DB 12; Length 5302;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCTATCTTTTGGATTATTAAGGGATTTTGGGGATTTTCGGCCATTGTTTAAAAAATG 60  
 DB 1872 CGGTCTATCTTTTGGATTATTAAGGGATTTTGGGGATTTTCGGCCATTGTTTAAAAAATG 60  
 QY 61 AGCTGATTTTACAAAAATTTAACGCGAATTAATTCGTGGA 101  
 DB 1932 AGCTGATTTTACAAAAATTTAACGCGAATTAATTCGTGGA 1972

RESULT 12  
 ADH11233  
 ID ADH11233 standard; DNA; 5425 BP.  
 XX  
 AC ADH11233;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Vertebrate UNC-53 protein homologue related nucleotide sequence.  
 XX  
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;  
 KW cell shape regulator; cell motility regulator; cell migration;  
 KW cell behaviour regulator; phenotype; signal transduction pathway;  
 KW signal transducing protein; signal integrator protein;  
 KW neuronal regeneration; revascularisation; wound healing;  
 KW chronic neurodegenerative disease; acute traumatic injury;  
 KW fibrotic disease; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9824810-A2.  
 XX  
 PD 11-JUN-1998.  
 XX  
 PF 03-DEC-1997; 97WO-EP006956.  
 XX  
 PR 04-DEC-1996; 96GB-00025283.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Plattreuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;  
 PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;  
 PI Geyse J, Bogaert TAOE;  
 XX  
 DR WPI; 1998-362411/31.  
 DR P-PSDB; ADH11234.  
 XX  
 PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.  
 PT promoting neuronal regeneration, treating chronic neuro-degenerative  
 PT diseases or acute traumatic injuries.  
 XX  
 PS Disclosure; Page 231-237; 479pp; English.  
 XX  
 CC The present invention describes a vertebrate protein homologue of an UNC-  
 CC 53 protein of Caenorhabditis elegans or a functional equivalent,  
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence  
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a  
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising  
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)

CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)  
 CC a compound identified as an enhancer or inhibitor of the regulation of  
 CC cell shape, motility, or the direction of cell migration for use as a  
 CC therapeutic; (7) a method for determination of whether a protein is an  
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or  
 CC motility or the direction of migration by contacting a host cell  
 CC expressing a homologue of UNC-53 and determining a change of phenotype;  
 CC (8) a method for identification of vertebrate homologues of C. elegans  
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to  
 CC a DNA library; and (9) a method for identification of a protein which is  
 CC active in the signal transduction pathway of a cell of which a vertebrate  
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract  
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an  
 CC antibody/homologue complex; and (iii) analysing such a complex to  
 CC identify any non-antibody protein bound to the complex. UNC-53 is a  
 CC signal transducing or signal integrator protein involved in controlling  
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate  
 CC homologues of UNC-53 can be used to promote neuronal regeneration,  
 CC revascularisation or wound healing, to treat chronic neurodegenerative  
 CC diseases or acute traumatic injuries or fibrotic diseases. The present  
 CC sequence is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 5425 BP; 1250 A; 1463 C; 1420 G; 1292 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 2; Length 5425;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCTATCTTTTGGATTATTAAGGGATTTTGGGGATTTTCGGCCATTGTTTAAAAAATG 60  
 DB 3664 CGGTCTATCTTTTGGATTATTAAGGGATTTTGGGGATTTTCGGCCATTGTTTAAAAAATG 3723  
 QY 61 AGCTGATTTTACAAAAATTTAACGCGAATTAATTCGTGGA 101  
 DB 3724 AGCTGATTTTACAAAAATTTAACGCGAATTAATTCGTGGA 3764  
 RESULT 13  
 ABN86685  
 ID ABN86685 standard; DNA; 5431 BP.  
 XX  
 AC ABN86685;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Nucleotide sequence of a pCDNA3 plasmid vector.  
 KW Major histocompatibility complex; MHC; antigen presenting cell; APC;  
 KW antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;  
 KW cancer; viral infection; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200261113-A2.  
 XX  
 PD 08-AUG-2002.  
 XX  
 PF 01-FEB-2002; 2002WO-US002598.  
 XX  
 PR 01-FEB-2001; 2001US-0265334P.  
 XX  
 PA (UJJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Wu T, Hung C;  
 XX  
 DR WPI; 2002-619261/66.  
 XX  
 PT Nucleic acid molecule encoding a fusion polypeptide that promotes  
 PT processing via the Major Histocompatibility Complex class I pathway  
 PT and/or promotes activity of an antigen presenting cell, useful as vaccine  
 PT for cancer and viral infections.  
 XX  
 PS Claim 24; Page 22-23; 127pp; English.

XX The invention relates to a new nucleic acid molecule (I) encoding a  
 CC fusion polypeptide useful as a vaccine composition. (I) comprises a first  
 CC nucleic acid sequence encoding a first polypeptide or peptide that  
 CC promotes processing via the Major Histocompatibility Complex (MHC) class  
 CC I pathway (MHC-I-PP) and/or promotes development or activity of an  
 CC antigen presenting cell (APC). The second nucleic acid sequence is linked  
 CC in frame to the first nucleic acid sequence or to a linker nucleic acid  
 CC sequence and encodes an antigenic polypeptide or peptide. The methods and  
 CC compositions of the present invention are useful as therapeutic vaccine  
 CC for cancer and for major viral infections, such as hepatoma and cervical  
 CC cancer, that cause morbidity and mortality. They can also be used in  
 CC treating animal diseases, such as equine herpesvirus, bovine viruses,  
 CC Marek's disease, retroviral and lentiviral diseases and rabies, in the  
 CC veterinary medicine context. The present sequence represents the  
 CC nucleotide sequence of a pcDNA3 plasmid vector  
 XX  
 SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 6; Length 5431;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 |||||  
 Db 1636 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1695  
 QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
 |||||  
 Db 1696 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1736  
 RESULT 14  
 ADE21866  
 ID ADE21866 standard; DNA; 5431 BP.  
 AC ADE21866;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Plasmid vector pcDNA3 nucleotide sequence SEQ ID NO:8.  
 XX  
 KW chimeric fusion; translocation; antigenic; cytostatic; immunotherapy;  
 KW gene therapy; cancer; tumour; gene; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003085085-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 04-APR-2003; 2003WO-US010235.  
 XX  
 PR 04-APR-2002; 2002US-00115440.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Wu T, Hung C;  
 XX  
 WPI; 2003-877027/81.  
 DR  
 XX  
 PT New nucleic acid encoding a chimeric fusion or fusion polypeptide  
 PT comprising a first domain with a translocation polypeptide, and a second  
 PT domain with an antigen having at least one antigenic peptide, useful for  
 PT treating cancer.  
 PT  
 PS Disclosure; SEQ ID NO 8; 68pp; English.  
 PS  
 XX The present invention describes a nucleic acid (I) encoding a chimeric  
 CC fusion or fusion polypeptide comprising a first domain with a  
 CC translocation polypeptide, and a second domain comprising an antigen  
 CC having at least one antigenic peptide. Also described: (i) an expression  
 CC vector comprising (I) operatively linked to a promoter and optionally, to

CC one or more regulatory elements that enhance expression of the nucleic  
 CC acid in a cell; (2) a particle comprising (I) or the expression vector;  
 CC (3) a cell that has been modified to comprise (I) or the expression  
 CC vector; (4) a chimeric polypeptide comprising a first domain with a  
 CC translocation polypeptide, and a second domain comprising an antigen  
 CC having at least one antigenic peptide; (5) a pharmaceutical composition  
 CC capable of inducing or enhancing an antigen specific immune response,  
 CC comprising (I), expression vector, particle, cell, cell of the particle,  
 CC or the chimeric polypeptide; and a carrier or excipient; (6) inducing or  
 CC enhancing an antigen specific immune response by administering the  
 CC composition described above; (7) increasing the number of CD8 + CTLs  
 CC specific for a selected desired antigen in a subject by administering the  
 CC composition described above; and (8) inhibiting the growth of a tumour in  
 CC a subject by administering the composition described above. (I) has  
 CC cytostatic activity, and can be used in immunotherapy, and gene therapy.  
 CC The nucleic acids (I), compositions and methods are useful for treating  
 CC cancer. The present sequence represents a plasmid vector nucleotide  
 CC sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 101; DB 10; Length 5431;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 |||||  
 Db 1636 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1695  
 QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101  
 |||||  
 Db 1696 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1736  
 RESULT 15  
 ADO05277  
 ID ADO05277 standard; DNA; 5431 BP.  
 XX  
 AC ADO05277;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE pcDNA3 plasmid vector.  
 XX  
 KW Translocation domain; bacterial toxin; exotoxin A domain II; ETA;  
 KW major histocompatibility complex; MHC class I; vaccine; immune response;  
 KW CD8+ cytotoxic T lymphocyte; CTL; tumour; E7 antigen; pcDNA3 plasmid; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN US2004086845-A1.  
 XX  
 PD 06-MAY-2004.  
 XX  
 PF 04-APR-2002; 2002US-00115440.  
 XX  
 PR 20-OCT-1999; 99US-00421608.  
 PR 09-FEB-2000; 2000US-00501097.  
 PR 20-OCT-2000; 2000WO-US041422.  
 PR 04-APR-2001; 2001US-0281003P.  
 XX  
 PA (WUTY/) WU T.  
 PA (HUNG/) HUNG C.  
 XX  
 PI Wu T, Hung C;  
 XX  
 WPI; 2004-356187/33.  
 DR  
 XX Novel chimeric polypeptide e.g., Pseudomonas aeruginosa exotoxin A domain  
 PT II/human papilloma virus-16 E7 peptide useful for inducing or enhancing  
 PT antigen specific immune response, or for inhibiting growth of tumor in  
 PT subject.  
 PT

PS Disclosure; SEQ ID NO 8; 48pp; English.

XX  
CC The invention relates to nucleic acid encoding a chimeric polypeptide  
CC comprising a translocation domain of a bacterial toxin and at least one  
CC antigenic peptide. The preferred translocation domain is domain II of  
CC pseudomonas aeruginosa exotoxin A (ETA(dII)) and the preferred antigen is  
CC human papilloma virus type 16 (HPV-16) E7 which is a model tumour  
CC antigen. The antigenic peptide comprises an epitope that binds to and is  
CC presented on the cell surface by major histocompatibility complex (MHC)  
CC class I proteins. The nucleic acid of the invention is useful as vaccine  
CC composition for enhancing antigen specific immune response, increasing  
CC the number of CD8+ cytotoxic T lymphocytes (CTLs) and for inhibiting the  
CC growth of a tumour. The present sequence is pcDNA3 plasmid vector used in  
CC the invention.

XX  
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5431;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTCATTCTTTTGATTATTAAGGATTTTCGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1636 CGGTCATTCTTTTGATTATTAAGGATTTTCGGGATTTTCGGCTATTGTTAAAAAATG 1695

QY 61 AGCTGATTTAACAAAAATTTAACGGGAATTAATTCGTGGA 101

Db 1696 AGCTGATTTAACAAAAATTTAACGGGAATTAATTCGTGGA 1736

Search completed: July 14, 2005, 07:01:25

Job time : 143.448 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_3674\_3774  
Perfect score: 101  
Sequence: 1 cggctctattctttgattta.....acgcgaattaattctgtgga 101

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88.4	87.5	827	5	BQ151093 NF044H03L
C 2	88.2	87.3	805	5	BQ145762 NF017H10G
C 3	88.2	87.3	813	5	BQ145733 NF017D06G
C 4	88	87.1	585	5	BQ145733 NF017D06G
C 5	87.4	86.5	750	5	BQ144787 NF104G07D
C 6	87.4	86.5	750	5	BQ154917 NF074C04I
C 7	87.4	86.5	764	5	BQ155881 NF085C08I
C 8	87.4	86.5	768	5	BQ155935 NF177H07P
C 9	87.4	86.5	783	5	BQ158754 NF070H02P
C 10	87.4	86.5	784	5	BQ159157 NF050D02P
C 11	87.4	86.5	784	5	BQ159352 NF114D02P
C 12	87.4	86.5	785	5	BQ156587 NF094D10I
C 13	87.4	86.5	796	5	BQ139350 NF014C09P
C 14	87.4	86.5	796	5	BQ153647 NF040H05I
C 15	87.4	86.5	798	5	BQ154182 NF056D06I
C 16	87.4	86.5	816	5	BQ144358 NF072C01D
C 17	87.4	86.5	820	5	BQ154225 NF056C02I
C 18	87.4	86.5	832	5	BQ146006 NF024A04G
C 19	87.2	86.3	389	4	B1938184 de35h05.Y
C 20	87.2	86.3	392	4	B1938175 de35g04.Y
C 21	87.2	86.3	395	6	CA620718 wlln.pk00
C 22	87.2	86.3	453	6	CA624892 wlln.pk01
C 23	87.2	86.3	472	4	BM276444 PfESToa8
C 24	87.2	86.3	565	6	CA623376 wlln.pk01

25	87.2	86.3	588	6	CA617091
26	86.8	85.9	137	6	CD282765
27	86.8	85.9	147	6	CD285004
c 28	86.8	85.9	154	6	CD288994
29	86.8	85.9	161	7	CN523248
30	86.8	85.9	164	6	CD282911
c 31	86.8	85.9	181	6	AQ014620
c 32	86.8	85.9	183	8	AQ014434
c 33	86.8	85.9	185	2	BF942500
c 34	86.8	85.9	188	8	B97287
35	86.8	85.9	193	6	CD281345
c 36	86.8	85.9	193	8	BZ666758
37	86.8	85.9	202	6	CD280039
c 38	86.8	85.9	204	8	B68786
39	86.8	85.9	205	6	CD280392
40	86.8	85.9	207	7	CN523315
41	86.8	85.9	208	7	CN521451
c 42	86.8	85.9	210	1	AJ409361
c 43	86.8	85.9	217	8	B60716
44	86.8	85.9	224	6	CD280393
45	86.8	85.9	235	6	CD280049

ALIGNMENTS

RESULT 1  
BQ151093/c  
LOCUS BQ151093 827 bp mRNA linear EST 24-APR-2002  
DEFINITION NF044H03L.F1029 Developing leaf Medicago truncatula cDNA clone  
ACCESSION BQ151093  
VERSION BQ151093.1 GI:20288152  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
REFERENCE 1 (bases 1 to 827)  
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,  
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.  
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
JOURNAL Medicago truncatula leaf library  
COMMENT Unpublished (2000)  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 224 6650  
Fax: 580 224 6692  
Email: gdmay@noble.org  
Insert Length: 827 Std Error: 0.00  
Plate: 044 row: H column: 03  
Seq primer: TCACACGAGAAACAGCTATGAC.  
Location/Qualifiers  
1. .827  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF044H03L.F"  
/tissue\_type="leaf"  
/dev\_stage="Pooled developmental"  
/clone\_lib="Developing leaf"  
/note="Vector: Lambda Zap; Contains a mixture of very  
young, developing, mature and senescing leaves."

Query Match 87.5%; Score 88.4; DB 5; Length 827;  
Best Local Similarity 98.9%; Pred. No. 1.1e-13;  
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db |||||||
483 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 424
|||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATT 90
Db |||||||
423 AGCTGATTAAACAAAAATTTAACGCGAATT 394
|||

RESULT 2
BQ145762/c
LOCUS
DEFINITION
  BQ145762 805 bp mRNA linear EST 24-APR-2002
  NF017H10GS1F1091 Germinating Seed Medicago truncatula cDNA clone
  NF017H10GS 5', mRNA sequence.
ACCESSION
  BQ145762
VERSION
  BQ145762.1 GI:20282821
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Medicago truncatula
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 805)
  Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
  Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula germinating seed library
  Unpublished (2001)
JOURNAL
  Contact: May GD
COMMENT
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 224 6650
  Fax: 580 224 6692
  Email: gdmay@noble.org
  Insert Length: 805 Std Error: 0.00
  Plate: 017 row: H column: 10
  Seq primer: TCACACGAGAAACAGCTATGAC.
  Location/Qualifiers
  1..805
  /organism="Medicago truncatula"
  /mol_type="mRNA"
  /db_xref="taxon:3880"
  /clone="NF017H10GS"
  /tissue_type="germinating seeds"
  /dev_stage="0, 1, 2 and 3 days after acid treatment."
  /clone_lib="Germinating Seed"
  /note="Vector: Lambda Zap; M. truncatula seeds were acid
  treated and placed on wet filter papers in petri dishes.
  Seeds were harvested at 0, 1, 2 and 3 days after acid
  treatment. cDNA was prepared from polyA+ enriched, pooled
  samples of equivalent amounts of total RNA from each time
  point. The cDNA was directionally ligated into the
  Uni-Zap XR vector (Stratagene) and packaged using the
  Gigapack III Gold packaging extracts. Phagemids
  containing cDNA inserts were in vivo excised from the
  recombinant Uni-Zap XR vector using ExAssist helper phage
  and the E. coli strain XLI-Blue MRF' (Stratagene).
  Excised plasmids were plated using SOLR cells."

FEATURES
  source
  1..805
  /organism="Medicago truncatula"
  /mol_type="mRNA"
  /db_xref="taxon:3880"
  /clone="NF017H10GS"
  /tissue_type="germinating seeds"
  /dev_stage="0, 1, 2 and 3 days after acid treatment."
  /clone_lib="Germinating Seed"
  /note="Vector: Lambda Zap; M. truncatula seeds were acid
  treated and placed on wet filter papers in petri dishes.
  Seeds were harvested at 0, 1, 2 and 3 days after acid
  treatment. cDNA was prepared from polyA+ enriched, pooled
  samples of equivalent amounts of total RNA from each time
  point. The cDNA was directionally ligated into the
  Uni-Zap XR vector (Stratagene) and packaged using the
  Gigapack III Gold packaging extracts. Phagemids
  containing cDNA inserts were in vivo excised from the
  recombinant Uni-Zap XR vector using ExAssist helper phage
  and the E. coli strain XLI-Blue MRF' (Stratagene).
  Excised plasmids were plated using SOLR cells."

ORIGIN
  Query Match 87.3%; Score 88.2; DB 5; Length 805;
  Best Local Similarity 92.1%; Pred. NO. 1.3e-13;
  Matches 93; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db |||||||
232 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 173
|||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTCTCTGTGA 101
Db |||||||
172 AGCTGATTAAACAAAAATTTAACGCGAATTCCTCGTGCGGA 132
|||

```

```

RESULT 3
BQ145733/c
LOCUS
DEFINITION
  BQ145733 813 bp mRNA linear EST 24-APR-2002
  NF017D06GS1F1057 Germinating Seed Medicago truncatula cDNA clone
  NF017D06GS 5', mRNA sequence.
ACCESSION
  BQ145733
VERSION
  BQ145733.1 GI:20282792
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Medicago truncatula
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 813)
  Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
  Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula germinating seed library
  Unpublished (2001)
JOURNAL
  Contact: May GD
COMMENT
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 224 6650
  Fax: 580 224 6692
  Email: gdmay@noble.org
  Insert Length: 813 Std Error: 0.00
  Plate: 017 row: D column: 06
  Seq primer: TCACACGAGAAACAGCTATGAC.
  Location/Qualifiers
  1..813
  /organism="Medicago truncatula"
  /mol_type="mRNA"
  /db_xref="taxon:3880"
  /clone="NF017D06GS"
  /tissue_type="germinating seeds"
  /dev_stage="0, 1, 2 and 3 days after acid treatment."
  /clone_lib="Germinating Seed"
  /note="Vector: Lambda Zap; M. truncatula seeds were acid
  treated and placed on wet filter papers in petri dishes.
  Seeds were harvested at 0, 1, 2 and 3 days after acid
  treatment. cDNA was prepared from polyA+ enriched, pooled
  samples of equivalent amounts of total RNA from each time
  point. The cDNA was directionally ligated into the
  Uni-Zap XR vector (Stratagene) and packaged using the
  Gigapack III Gold packaging extracts. Phagemids
  containing cDNA inserts were in vivo excised from the
  recombinant Uni-Zap XR vector using ExAssist helper phage
  and the E. coli strain XLI-Blue MRF' (Stratagene).
  Excised plasmids were plated using SOLR cells."

ORIGIN
  Query Match 87.3%; Score 88.2; DB 5; Length 813;
  Best Local Similarity 92.1%; Pred. NO. 1.3e-13;
  Matches 93; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db |||||||
231 CGGCTATTCTTTGATTATAAGGATTTGGGATTTCCGCCCTATTGGTTAAAAAATG 172
|||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATTCCTGTGA 101
Db |||||||
171 AGCTGATTAAACAAAAATTTAACGCGAATTCCTCGTGCGGA 131
|||

RESULT 4
BU998405
LOCUS
DEFINITION
  BU998405 585 bp mRNA linear EST 23-OCT-2002
  H110P06r HI Hordeum vulgare subsp. vulgare cDNA clone H110P06
  5-PRIME, mRNA sequence.

```

```

ACCESSION   BU998405
VERSION     BU998405.1  GI:24275388
KEYWORDS
SOURCE
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Poaceae; Triticeae; Hordeum.
            1 (bases 1 to 585)
REFERENCE   Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
AUTHORS
TITLE       EST sequencing and analysis in barley (2002)
JOURNAL
COMMENT     Contact: Stein Nils
            Molecular Markers Group, Department Genbank
            Institute of Plant Genetics and Crop Plant Research (IPK)
            Corrensstr. 3, 06466, Gatersleben, Germany
            Tel: 039482-5522
            Fax: 039482-5595
            Email: stein@ipk-gatersleben.de
            Insert Length: 585 Std Error: 0.00
            Plate: 10 row: P column: 6
            Seq primer: M13rev.
FEATURES   Location/Qualifiers
            source
            1..585
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="barke"
            /sub_species="vulgare"
            /db_xref="GABI:252269"
            /db_xref="taxon:112509"
            /clone="H110P06"
            /tissue_type="female inflorescences"
            /dev_stage="female inflorescences (approx. 3 mm in size)"
            /lab_host="XL10-Gold"
            /clone_lib="H1"
            /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
            cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
            artefact caused by the kit, in most cases the EcoRI site
            is NOT present, as well as the EcoRI adapter used for
            cloning. To excise the insert, restriction sites upstream
            EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
            due to the cloning system used Blue/white selection for
            recombinants is not 100% reliable."
ORIGIN
Query Match      87.1%; Score 88; DB 5; Length 585;
Best Local Similarity 94.8%; Pred. No. 1.4e-13;
Matches 91; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db 54 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 113
Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTT 96
Db 114 AGCTGATTAAACAAAAATTTAACGCGAATTTTIT 149
RESULT 5
BQ144787/c
LOCUS       BQ144787 750 bp mRNA linear EST 24-APR-2002
DEFINITION  NF104G07D7IF1054 Drought Medicago truncatula cDNA clone NF104G07D
            5', mRNA sequence.
ACCESSION   BQ144787
VERSION     BQ144787.1  GI:20281846
KEYWORDS
SOURCE
  ORGANISM  Medicago truncatula (barrel medic)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
            1 (bases 1 to 750)
REFERENCE   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
AUTHORS
TITLE       Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
JOURNAL
COMMENT     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula drought library
            Unpublished (2001)
            Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 224 6650
            Fax: 580 224 6692
            Email: gdmay@noble.org
            Insert Length: 750 Std Error: 0.00
            Plate: 074 row: C column: 04
            Seq primer: TCACACAGGAACACGTATGAC.
FEATURES   Location/Qualifiers
            source
            1..750
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF104G07DT"
            /tissue_type="Plantlets"
            /dev_stage="Pooled timepoints"
            /clone_lib="Drought"
            /note="Vector: Lambda Zap; Contains a mixture of entire
            plantlets harvested in a series of days-post-watering
            timepoints."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db 192 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 133
Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTG 97
Db 132 AGCTGATTAAACAAAAATTTAACGCGAATTCCTGCAG 96
RESULT 6
BQ154917/c
LOCUS       BQ154917 750 bp mRNA linear EST 24-APR-2002
DEFINITION  NF074C04IR 5', mRNA sequence.
ACCESSION   BQ154917
VERSION     BQ154917.1  GI:20291976
KEYWORDS
SOURCE
  ORGANISM  Medicago truncatula (barrel medic)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
            1 (bases 1 to 750)
REFERENCE   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
AUTHORS
TITLE       Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
JOURNAL
COMMENT     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula irradiated library
            Unpublished (2001)
            Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 224 6650
            Fax: 580 224 6692
            Email: gdmay@noble.org
            Insert Length: 750 Std Error: 0.00
            Plate: 074 row: C column: 04
            Seq primer: TCACACAGGAACACGTATGAC.
FEATURES   Location/Qualifiers
            source
            1..750
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF104G07DT"
            /tissue_type="Plantlets"
            /dev_stage="Pooled timepoints"
            /clone_lib="Drought"
            /note="Vector: Lambda Zap; Contains a mixture of entire
            plantlets harvested in a series of days-post-watering
            timepoints."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
Db 192 CGGTCATCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 133
Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTG 97
Db 132 AGCTGATTAAACAAAAATTTAACGCGAATTCCTGCAG 96

```

```

source
1. .750
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF074C041R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAassist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."

ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTCTTTGATTATAGGGATTTCGGGCTATTGGCTATTGGTTAAAAATG 60
    |||
Db 192 CGGTCATCTCTTTGATTATAGGGATTTCGGGCTATTGGCTATTGGTTAAAAATG 133
    |||

Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAAATCTG 97
    |||
Db 132 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 96
    |||

RESULT 7
BQ155881/c
LOCUS
DEFINITION
  NF085C08IR1F1066 Irradiated Medicago truncatula cDNA clone
  NF085C08IR 5', mRNA sequence.
ACCESSION
  BQ155881
VERSION
  BQ155881.1 GI:20292940
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Medicago truncatula
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 764)
  Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
  Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula irradiated library
  Unpublished (2001)
  Contact: May GD
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 224 6650
  Fax: 580 224 6692
  Email: gdmay@noble.org
  Insert Length: 764 Std Error: 0.00
  Plate: 085 row: C column: 08
  Seq primer: TCACACGGAACAGCTATGAC.
  Location/Qualifiers
    1..764
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /db_xref="taxon:3880"
    /clone="NF085C08IR"

FEATURES
source
1. .764
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF074C041R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAassist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."

ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTCTTTGATTATAGGGATTTCGGGCTATTGGCTATTGGTTAAAAATG 60
    |||
Db 191 CGGTCATCTCTTTGATTATAGGGATTTCGGGCTATTGGCTATTGGTTAAAAATG 132
    |||

Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAAATCTG 97
    |||
Db 131 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 95
    |||

RESULT 8
BQ159395/c
LOCUS
DEFINITION
  NF117H07PL1F1064 Phosphate starved leaf Medicago truncatula cDNA
  clone NF117H07PL 5', mRNA sequence.
ACCESSION
  BQ159395
VERSION
  BQ159395.1 GI:20296452
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Medicago truncatula
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 768)
  Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
  Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula phosphate-starved leaf library
  Unpublished (2000)
  Contact: Harrison MJ
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 221 7325
  Fax: 580 221 7380
  Email: mjharrison@noble.org
  Insert Length: 768 Std Error: 0.00
  Plate: 117 row: H column: 07
  Seq primer: TCACACGGAACAGCTATGAC.
  Location/Qualifiers
    1..768
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /db_xref="taxon:3880"
    /clone="NF117H07PL"
    /tissue_type="leaf"
    /dev_stage="trifoliolate"
    /clone_lib="Phosphate starved leaf"
    /note="Vector: Lambda Zap; At the trifoliolate stage, M.
    truncatula plants were transplanted to phosphate-free sand

```



and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."

## ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 768;  
 Best Local Similarity 93.8%; Pred. No. 2.1e-13;  
 Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 193 CGGCTATTCTTTGATTATTAAGGATTTGGCCATTTGGCTATTGGTTAAAAAATG 134

Qy 61 AGCTGATTACAAAAATTTAACGCGAATTAATTCG 97  
 Db 133 AGCTGATTACAAAAATTTAACGCGAATTCCTGCAG 97

## RESULT 9

BQ158754/c  
 LOCUS BQ158754 783 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF070H02PL1F1026 Phosphate starved leaf Medicago truncatula cDNA  
 clone NF070H02PL 5', mRNA sequence.  
 ACCESSION BQ158754  
 VERSION BQ158754.1 GI:20295811  
 KEYWORDS EST.  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 783)  
 AUTHORS Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.  
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula phosphate-starved leaf library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Harrison MJ  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7325  
 Fax: 580 221 7380  
 Email: mjharrison@noble.org  
 Insert Length: 783 Std Error: 0.00  
 Plate: 070 row: H column: 02  
 Seq primer: TCACACGGAACAGCTATGAC.  
 Location/Qualifiers  
 1..783  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF070H02PL"  
 /tissue\_type="leaf"  
 /dev\_stage="trifoliolate"  
 /clone\_lib="Phosphate starved leaf"  
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."

## FEATURES

source

## ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 783;  
 Best Local Similarity 93.8%; Pred. No. 2.1e-13;  
 Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 192 CGGCTATTCTTTGATTATTAAGGATTTGGCCATTTGGCTATTGGTTAAAAAATG 133

Qy 61 AGCTGATTACAAAAATTTAACGCGAATTAATTCG 97  
 Db 132 AGCTGATTACAAAAATTTAACGCGAATTCCTGCAG 96

RESULT 10  
 BQ159157/c

LOCUS BQ159157 784 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF050D02PL1F1014 Phosphate starved leaf Medicago truncatula cDNA  
 clone NF050D02PL 5', mRNA sequence.

ACCESSION BQ159157  
 VERSION BQ159157.1 GI:20296214  
 KEYWORDS EST.  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 784)

AUTHORS Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.  
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula phosphate-starved leaf library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Harrison MJ  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7325  
 Fax: 580 221 7380  
 Email: mjharrison@noble.org  
 Insert Length: 784 Std Error: 0.00  
 Plate: 050 row: D column: 02  
 Seq primer: TCACACGGAACAGCTATGAC.  
 Location/Qualifiers  
 1..784  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF050D02PL"  
 /tissue\_type="leaf"  
 /dev\_stage="trifoliolate"  
 /clone\_lib="Phosphate starved leaf"  
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."

## FEATURES

source

## ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 784;  
 Best Local Similarity 93.8%; Pred. No. 2.1e-13;  
 Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60  
 Db 192 CGGCTATTCTTTGATTATTAAGGATTTGGCCATTTGGCTATTGGTTAAAAAATG 133

Qy 61 AGCTGATTACAAAAATTTAACGCGAATTAATTCG 97  
 Db 132 AGCTGATTACAAAAATTTAACGCGAATTCCTGCAG 96

RESULT 11  
 BQ159352/c

LOCUS BQ159352 784 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF114D02PL1F1026 Phosphate starved leaf Medicago truncatula cDNA  
 clone NF114D02PL 5', mRNA sequence.

ACCESSION BQ159352  
 VERSION BQ159352.1 GI:20296409

```

KEYWORDS      EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
REFERENCE      1 (bases 1 to 784)
AUTHORS       Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
               Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
TITLE         Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula phosphate-starved leaf library
JOURNAL       Unpublished (2000)
COMMENT       Contact: Harrison MJ
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73402, USA
               Tel: 580 221 7325
               Fax: 580 221 7380
               Email: mharrison@noble.org
               Insert Length: 784 Std Error: 0.00
               Plate: 114 row: D column: 02
               Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES      Location/Qualifiers
               source
               1..784
               /organism="Medicago truncatula"
               /mol_type="mRNA"
               /db_xref="taxon:3880"
               /clone="NF114D02PL"
               /tissue_type="leaf"
               /dev_stage="trifoliolate"
               /clone_lib="Phosphate starved leaf"
               /note="Vector: Lambda Zap; At the trifoliolate stage, M.
               truncatula plants were transplanted to phosphate-free sand
               and grown for a further 30 days. During this 30 day
               period, the plants were fertilized twice weekly with 1/2
               Hoaglands solution containing only 20uM potassium
               phosphate. RNA was prepared from above ground tissues."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 784;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATATAAGGGAATTTGGGGAATTCGGCCTATTGGTTAAAAAATG 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 194 CGGTCATCTTTGATTATATAAGGGAATTTGGGGAATTCGGCCTATTGGTTAAAAAATG 135

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAATTCG 97
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 134 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 98

RESULT 12
BQ156587/c
LOCUS          BQ156587.1 785 bp mRNA linear EST 24-APR-2002
DEFINITION    NF094D10IR1F090 Irradiated Medicago truncatula cDNA clone
               NF094D10IR 5', mRNA sequence.
ACCESSION     BQ156587
VERSION       BQ156587.1 GI:20293646
KEYWORDS      EST.
ORGANISM      Medicago truncatula (barrel medic)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
REFERENCE      1 (bases 1 to 785)
AUTHORS       Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
               Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE         Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula irradiated library
JOURNAL       Unpublished (2001)

KEYWORDS      EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
REFERENCE      1 (bases 1 to 796)
AUTHORS       Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
               Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
               May,G.D. and Paiva,N.L.
TITLE         Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula Phoma-infected library
JOURNAL       Unpublished (2002)
COMMENT       Contact: Paiva NL
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 785 Std Error: 0.00
Plate: 094 row: D column: 10
Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES      Location/Qualifiers
               source
               1..785
               /organism="Medicago truncatula"
               /mol_type="mRNA"
               /db_xref="taxon:3880"
               /clone="NF094D10IR"
               /tissue_type="seedlings"
               /dev_stage="seedling"
               /clone_lib="Irradiated"
               /note="Vector: Lambda Zap; Seedlings were exposed either
               to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
               Gamma-irradiated samples were harvested at 6, 12, 24 and
               48 hours after treatment. UV-irradiated samples were
               harvested 24 hours post-treatment. cDNA was prepared from
               polyA+ enriched, pooled samples of equivalent amounts of
               total RNA from each sample. The cDNA was directionally
               ligated into the Uni-Zap XR vector (Stratagene) and
               packaged using the Gigapack III Gold packaging extracts.
               Phagemids containing cDNA inserts were in vivo excised
               from the recombinant Uni-Zap XR vector using ExAssist
               helper phage and the E. coli strain XL1-Blue MRF'
               (Stratagene). Excised plasmids were plated using SOLR
               cells."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 785;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATATAAGGGAATTTGGGGAATTCGGCCTATTGGTTAAAAAATG 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191 CGGTCATCTTTGATTATATAAGGGAATTTGGGGAATTCGGCCTATTGGTTAAAAAATG 132

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAATTCG 97
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 95

RESULT 13
BQ139350/c
LOCUS          BQ139350 796 bp mRNA linear EST 26-APR-2002
DEFINITION    NF014C09PH1F1070 Phoma-infected Medicago truncatula cDNA clone
               NF014C09PH 5', mRNA sequence.
ACCESSION     BQ139350
VERSION       BQ139350.1 GI:20275476
KEYWORDS      EST.
ORGANISM      Medicago truncatula (barrel medic)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
REFERENCE      1 (bases 1 to 796)
AUTHORS       Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
               Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
               May,G.D. and Paiva,N.L.
TITLE         Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula Phoma-infected library
JOURNAL       Unpublished (2002)
COMMENT       Contact: Paiva NL
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73402, USA

```

```

Tel: 580 221 7317
Fax: 580 221 7380
Email: nlpaiva@noble.org
Insert Length: 796 Std Error: 0.00
Plate: 014 row: C column: 09
Seq primer: TCACACGGAACACGCTATGAC.
Location/Qualifiers
FEATURES
    source
    1..796
        /organism="Medicago truncatula"
        /mol_type="mRNA"
        /db_xref="taxon:3880"
        /clone="NF014C09PH"
        /tissue_type="leaf"
        /dev_stage="Pathogen-induced, young trifoliolate"
        /note="Vector: pBluescript SK(-); Young trifoliolate leaves of Medicago truncatula were excised and dip-inoculated in humid dishes. Pools of leaves were harvested at 0, 15, and 30 minutes and 1, 2, 3, 6, 14, 24, 48, 72, and 96, hours, and used to prepare total RNA. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."
ORIGIN
    Query Match      86.5%; Score 87.4; DB 5; Length 796;
    Best Local Similarity 93.8%; Pred. No. 2.1e-13;
    Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 60
Db 191 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 132
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTG 97
Db 131 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 95
RESULT 14
BQ153647/c
LOCUS
DEFINITION
    NF040H05IR1F1047 Irradiated Medicago truncatula cDNA clone
ACCESSION
    BQ153647
VERSION
    BQ153647.1 GI:20290706
KEYWORDS
    EST.
SOURCE
    Medicago truncatula (barrel medic)
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
    Medicago.
REFERENCE
    1 (bases 1 to 796)
    Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
    Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
    Expressed Sequence Tags from the Samuel Roberts Noble Foundation
    Medicago truncatula irradiated library
    Unpublished (2001)
    Contact: May GD
    Plant Biology Division
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 796 Std Error: 0.00
JOURNAL
COMMENT
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 796 Std Error: 0.00
FEATURES
    source
    1..796
        /organism="Medicago truncatula"
        /mol_type="mRNA"
        /db_xref="taxon:3880"
        /clone="NF014C09PH"
        /tissue_type="leaf"
        /dev_stage="Pathogen-induced, young trifoliolate"
        /note="Vector: pBluescript SK(-); Young trifoliolate leaves of Medicago truncatula were excised and dip-inoculated in humid dishes. Pools of leaves were harvested at 0, 15, and 30 minutes and 1, 2, 3, 6, 14, 24, 48, 72, and 96, hours, and used to prepare total RNA. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."
ORIGIN
    Query Match      86.5%; Score 87.4; DB 5; Length 796;
    Best Local Similarity 93.8%; Pred. No. 2.1e-13;
    Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 60
Db 191 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 132
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTG 97
Db 131 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 95
RESULT 15
BQ154182/c
LOCUS
DEFINITION
    NF056D06IR1F1057 Irradiated Medicago truncatula cDNA clone
ACCESSION
    BQ154182
VERSION
    BQ154182.1 GI:20291241
KEYWORDS
    EST.
SOURCE
    Medicago truncatula (barrel medic)
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
    Medicago.
REFERENCE
    1 (bases 1 to 798)
    Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
    Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
    Expressed Sequence Tags from the Samuel Roberts Noble Foundation
    Medicago truncatula irradiated library
    Unpublished (2001)
    Contact: May GD
    Plant Biology Division
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 798 Std Error: 0.00
JOURNAL
COMMENT
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 798 Std Error: 0.00
FEATURES
    source
    1..798
        /organism="Medicago truncatula"
        /mol_type="mRNA"
        /db_xref="taxon:3880"
        /clone="NF040H05IR"
        /tissue_type="seedling"
        /dev_stage="seedling"
        /clone_lib="Irradiated"
        /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."
ORIGIN
    Query Match      86.5%; Score 87.4; DB 5; Length 796;
    Best Local Similarity 93.8%; Pred. No. 2.1e-13;
    Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 60
Db 193 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 134
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTG 97
Db 133 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 97
RESULT 15
BQ154182/c
LOCUS
DEFINITION
    NF056D06IR1F1057 Irradiated Medicago truncatula cDNA clone
ACCESSION
    BQ154182
VERSION
    BQ154182.1 GI:20291241
KEYWORDS
    EST.
SOURCE
    Medicago truncatula (barrel medic)
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
    Medicago.
REFERENCE
    1 (bases 1 to 798)
    Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
    Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
    Expressed Sequence Tags from the Samuel Roberts Noble Foundation
    Medicago truncatula irradiated library
    Unpublished (2001)
    Contact: May GD
    Plant Biology Division
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 798 Std Error: 0.00
JOURNAL
COMMENT
    The Samuel Roberts Noble Foundation
    2510 Sam Noble Parkway, Ardmore, OK 73402, USA
    Tel: 580 224 6650
    Fax: 580 224 6692
    Email: gdmay@noble.org
    Insert Length: 798 Std Error: 0.00
FEATURES
    source
    1..798
        /organism="Medicago truncatula"
        /mol_type="mRNA"
        /db_xref="taxon:3880"
        /clone="NF040H05IR"
        /tissue_type="seedling"
        /dev_stage="seedling"
        /clone_lib="Irradiated"
        /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."
ORIGIN
    Query Match      86.5%; Score 87.4; DB 5; Length 796;
    Best Local Similarity 93.8%; Pred. No. 2.1e-13;
    Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 60
Db 193 CGGTCTATCTTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAATG 134
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATCTG 97
Db 133 AGCTGATTTAACAAAAATTTAACGCGAATTCCTGCAG 97

```

```

/mol_type="mRNA"
/db_xref="taxon:380"
/clone="NF056D061R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/note="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAssist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."

```

ORIGIN

```

Query Match      86.5%; Score 87.4; DB 5; Length 798;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCTATCTTTGATTATAAGGGATTTTGGGGATTTTCGGCTATTGGTTAAAAAATG 60
    |||||||
Db 191 CGGTCTATCTTTGATTATAAGGGATTTTGGGGATTTTCGGCTATTGGTTAAAAAATG 132
    |||||||

Qy 61 AGCTGATTTAACAAAAATTTAAACGGAATTAAATTCG 97
    |||||||
Db 131 AGCTGATTTAACAAAAATTTAAACGGAATTCTCTGAG 95
    |||||||

```

Search completed: July 14, 2005, 23:22:30  
Job time : 967.667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds  
(without alignments)  
6468.225 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_7369\_7469  
Perfect score: 101  
Sequence: 1 agggattattgtctcatgagc.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	6 AR356490	AR356490 Sequence
C 2	101	100.0	142	6 AR538046	AR538046 Sequence
C 3	101	100.0	228	6 E00019	E00019 DNA coding
C 4	101	100.0	240	1 PM00END0	M10199 Plasmid pMM
C 5	101	100.0	251	6 E00018	E00018 DNA coding
C 6	101	100.0	251	6 I01644	I01644 Sequence 1
C 7	101	100.0	344	11 HUMUT5345	L18624 Human chrom
C 8	101	100.0	400	6 B0195256	B0195256 Nucleotid
C 9	101	100.0	456	6 E00892	E00892 Synthetic D
C 10	101	100.0	456	6 E01156	E01156 DNA fragmen
C 11	101	100.0	456	6 E01274	E01274 DNA encodin
C 12	101	100.0	456	6 E01302	E01302 DNA encodin
C 13	101	100.0	466	6 AX260098	AX260098 Sequence
C 14	101	100.0	573	6 AX260150	AX260150 Sequence
C 15	101	100.0	693	6 A43586	A43586 Sequence 11
C 16	101	100.0	698	6 AR116755	AR116755 Sequence
C 17	101	100.0	998	6 AY559171	AY559171 Pseudomon
C 18	101	100.0	1011	1 SMTMAOQE	X97254 S.marcescen
C 19	101	100.0	1012	2 C8C11F10	Z92776 Caenorhabdi

20	101	100.0	1014	4 CFAJ4121	AJ224121 Canis fam
C 21	101	100.0	1027	1 AY589493	AY589493 Escherich
C 22	101	100.0	1040	1 AY538698	AY538698 Serratia
C 23	101	100.0	1040	1 AY538700	AY538700 Serratia
C 24	101	100.0	1040	1 AY538701	AY538701 Serratia
C 25	101	100.0	1040	1 AY538702	AY538702 Serratia
C 26	101	100.0	1041	1 AY538699	AY538699 Serratia
C 27	101	100.0	1042	1 AY394610	AY394610 Klebsiell
C 28	101	100.0	1042	1 ECO308558	ECO308558 Escherich
C 29	101	100.0	1044	1 AY392531	AY392531 Streptoco
C 30	101	100.0	1044	1 AY452662	AY452662 Streptoco
C 31	101	100.0	1054	1 AF104441	AF104441 Klebsiell
C 32	101	100.0	1054	1 AF104442	AF104442 Escherich
C 33	101	100.0	1058	6 I03356	I03356 Sequence 4
C 34	101	100.0	1064	1 AY628199	AY628199 Escherich
C 35	101	100.0	1069	1 AF535127	AF535127 Klebsiell
C 36	101	100.0	1069	1 AY243512	AY243512 Klebsiell
C 37	101	100.0	1071	1 AY628175	AY628175 Escherich
C 38	101	100.0	1072	1 AY101764	AY101764 Klebsiell
C 39	101	100.0	1073	6 AR371489	AR371489 Sequence
C 40	101	100.0	1073	6 AX195443	AX195443 Sequence
C 41	101	100.0	1075	1 AY729027	AY729027 Proteus m
C 42	101	100.0	1075	1 PATN1PN2	X54605 Pseudomonas
C 43	101	100.0	1075	1 PATN2PN1B	X54607 Pseudomonas
C 44	101	100.0	1075	1 PATN3PN1A	X54604 Pseudomonas
C 45	101	100.0	1080	1 AF027199	AF027199 Klebsiell

ALIGNMENTS

RESULT 1	AR356490/c	142 bp	DNA	linear	PAT 17-AUG-2003
LOCUS	Sequence 2608 from patent US 6593114.				
DEFINITION	AR356490				
ACCESSION	AR356490.1				
VERSION	GI:33762574				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 142)				
AUTHORS	Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.				
TITLE	Staphylococcus aureus polynucleotides and sequences				
JOURNAL	Patent: US 6593114-A 2608 15-JUL-2003;				
FEATURES	Location/Qualifiers				
source	1..142				
	/organism="unknown"				
	/mol_type="genomic DNA"				

ORIGIN	Query Match	100.0%;	Score 101;	DB 6;	Length 142;
	Best Local Similarity	100.0%;	Pred. No. 8.7e-20;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGGTTATTGTCATGACGGGATACATATTGTAATGCTATTAGCAAAATAAACAATAG	60		
Db	107	AGGGTTATTGTCATGACGGGATACATATTGTAATGCTATTAGCAAAATAAACAATAG	48		
Qy	61	GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC	101		
Db	47	GGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTC	7		
RESULT 2	AR538046/c	142 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	Sequence 2608 from patent US 6737248.				
DEFINITION	AR538046				
ACCESSION	AR538046.1				
VERSION	GI:53929263				
KEYWORDS	Unknown.				
SOURCE	Unknown.				

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
source 1..142
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 48

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 101
Db 47 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 7

RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 228)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
source 1..228
Location/Qualifiers
1..228
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 116

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 101

```

```

Db 115 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 75

RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES Location/Qualifiers
source 1..240
/organism="Plasmid pMM110"
/mol_type="genomic DNA"
/db_xref="taxon:2599"
/plasmid="Plasmid pMM110"
ORIGIN Unreported.

Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 92

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 101
Db 91 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAGTC 51

RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 251)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;

```

```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES             Location/Qualifiers
     source          1..251
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
ORIGIN
Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 60
   |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 116
   |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
   |||||||
Db 115 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 75

RESULT 6
I01644/c
LOCUS I01644 251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES             Location/Qualifiers
     source          1..251
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 60
   |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 116
   |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
   |||||||
Db 115 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 75

RESULT 7
HUMUT5345
LOCUS HUMUT5345 344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker; microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

```

```

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAACAGGAGCGCAAAATGC
Primer B: TTCGGGAAATGTCCCGGAACC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2
FEATURES             Location/Qualifiers
     source          1..344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="8"
                     36..224
                     /standard_name="STS UT5345"
                     36..60
                     complement(202..224)
     STS
     primer_bind
     primer_bind
ORIGIN
Query Match          100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 60
   |||||||
Db 141 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAAG 200
   |||||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
   |||||||
Db 201 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 241

RESULT 8
BD195256/c
LOCUS BD195256 400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
1 (bases 1 to 400)
AUTHORS Dillion,P.J., Choi,G.H. and Welch,R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT
OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
PATRICK J DILLON,GIL H CHOI,RODNEY A WELCH
PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.
FEATURES
  source
    Location/Qualifiers
    1..400
    /organism='unidentified'
    /mol_type='genomic DNA'
    /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 165 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 106
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 105 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clones=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FH of beta-lactamase
FH RBS 200..203
FH CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FEATURES
  source
    Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'
```

```

ORIGIN
/db_xref='taxon:32630'
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
CC Key Location/Qualifiers
FH promoter 125..170
FH /note='beta lactamase promoter' FT RBS
FH CDS 200..204
FH /product='beta urogastrone'
FH sig_peptide 209..277
FH mat_peptide 278..436
FH /product='beta urogastrone'.
FEATURES
  source
    Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'
    /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
```



```
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaihara,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIDE, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FEATURES
source Location/Qualifiers
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source Location/Qualifiers
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION E01302.1 GI:2169561
KEYWORDS JP 1987190083-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Koiide,T. and Momota,Y.
TITLE POLYPEPTIDE-EXPRESSION VECTOR, HOST TRANSFORMED WITH SAID VECTOR
AND PRODUCTION OF POLYPEPTIDE USING SAID HOST
JOURNAL Patent: JP 1987190083-A 1 20-AUG-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OS Homo sapiens
PN JP 1987190083-A/1
PD 20-AUG-1987
PF 14-FEB-1986 JP 1986031415
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, KOIDE TAKAO,
MOMOTA YUTAKA
PC C12N15/00,C07H21/04,C12N1/00,C12P21/02,(C12N1/00,C12R1:19), PC
(C12P21/02);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FEATURES
source Location/Qualifiers
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
```

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 221
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 220 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 180
    |||||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION        AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM        Drosophila melanogaster (fruit fly)
                Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS        Deak, P., Glover, D.M. and Midgley, C.
TITLE          Cell cycle progression proteins
JOURNAL        Patent: WO 0172774-A 112 04-OCT-2001;
                Cyclacel Limited (GB)
FEATURES
source
1..573
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 296
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 295 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 255
    |||||||

RESULT 15
A43586
LOCUS
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION        A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM        Cuphea lanceolata
                Cuphea lanceolata
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; Myrtales; Lythraceae; Cuphea.
                1 (bases 1 to 693)
                Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
                Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
                Schulte, W., Voetz, M., Walek, J. and Schell, J.
REFERENCE
AUTHORS
PROMOTERS
TITLE          Patent: WO 9507357-A 11 16-MAR-1995;
JOURNAL        MAX PLANCK GESELLSCHAFT (DE)
COMMENT        Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 651
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 652 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 692
    |||||||

Search completed: July 14, 2005, 14:03:24
Job time : 759.618 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds  
(without alignments)  
4198.742 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_7369\_7469  
Perfect score: 101  
Sequence: 1 agggttattgtctatgagc.....gaaagtgccacctgacgtc 101

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	142	2 AAV76919	AAV76919 Staphyloc
C 2	101	100.0	228	1 AAN10032	Aan10032 Sequence
C 3	101	100.0	251	1 AAN10031	Aan10031 Sequence
C 4	101	100.0	400	2 AAV31229	AAV31229 E. coli J
C 5	101	100.0	456	1 AAN60624	Aan60624 Plasmid p
C 6	101	100.0	456	1 AAN71080	Aan71080 Sequence
C 7	101	100.0	456	1 AAN70833	Aan70833 Beta-urog
C 8	101	100.0	456	1 AAN81765	Aan81765 Sequence
C 9	101	100.0	466	6 ABA90413	ABA90413 Drosophil
C 10	101	100.0	487	2 AAX21173	Aax21173 Polynucle
C 11	101	100.0	535	2 AAX21149	Aax21149 Polynucle
C 12	101	100.0	573	6 ABA90456	ABA90456 Drosophil
C 13	101	100.0	605	12 ADH58311	Adh58311 Electroph
C 14	101	100.0	776	4 AAS30560	Aas30560 DNA encod
C 15	101	100.0	776	4 AAS27819	Aas27819 DNA encod
C 16	101	100.0	776	4 ABK42984	Abk42984 Genomic s
C 17	101	100.0	776	4 AAL07344	Aal07344 Human rep
C 18	101	100.0	776	4 AAL03229	Aal03229 Human rep
C 19	101	100.0	776	4 AAL06588	Aal06588 Human rep
C 20	101	100.0	776	4 AAL07340	Aal07340 Human rep

C 21	101	100.0	776	5 ABA14573	Abal14573 Human ner
C 22	101	100.0	776	5 AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8 ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8 ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8 ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9 ADB91869	Adb91869 Human sec
C 27	101	100.0	776	9 ADB61140	Adb61140 Connectiv
C 28	101	100.0	776	10 ADB94622	Adb94622 Novel hum
C 29	101	100.0	776	10 ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10 ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12 ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4 AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4 AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4 ABK42983	Abk42983 Genomic s
C 35	101	100.0	845	4 AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4 AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4 AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4 AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4 AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4 AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4 AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4 AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5 ABA14572	Abal14572 Human ner
C 44	101	100.0	845	5 AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9 ADB61139	Adb61139 Connectiv

ALIGNMENTS

RESULT 1  
AAV76919/c  
ID AAV76919 standard; DNA; 142 BP.  
XX  
AC AAV76919;  
XX  
DT 16-MAR-1999 (first entry)  
XX  
DE Staphylococcus aureus contig SEQ ID #2608.

KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.  
XX Staphylococcus aureus.

OS Staphylococcus aureus.

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
XX stored on computer readable medium and used in the production of anti-  
XX S.aureus vaccines.

XX Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
XX of the invention. The DNA sequences are recorded on a computer readable  
XX medium, preferably selected from a floppy or hard disk, random access  
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
XX the S.aureus DNA sequences allows putative functions to be assigned so  
XX that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can  
 CC be used in a vaccine composition against *S. aureus* infection. The  
 CC polypeptides can also be used in a kit for the immunodetection of  
 CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,  
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
 CC for recombinant production of the polypeptides. The new DNA sequences  
 CC (and their fragments) are useful as primers or probes for isolating  
 CC homologues of any of the *S. aureus* DNA sequences contained on the computer  
 CC readable medium

XX SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60  
 |||||  
 Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 48  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
 |||||  
 Db 47 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 7  
 |||||

## RESULT 2

AAAN10032/c  
 ID AAAN10032 standard; DNA; 228 BP.

XX AC AAAN10032;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;  
 XX insulin; ds.

XX OS Escherichia coli.

Key	Location/Qualifiers
FT misc_feature	1..4
FT	/tag= a
FT	/label= sticky end
FT misc_feature	225..228
FT	/tag= b
FT	/label= sticky end

XX EP38182-A.

XX PN 21-OCT-1981.

XX PF 09-APR-1981; 81EP-00301561.

XX PR 11-APR-1980; 80US-00139225.

XX PA (HARD ) HARVARD COLLEGE.

XX PI Gilbert W, Talmadge K;

XX DR WPI; 1981-80125D/44.

XX DR P-PSDB; AAP10039.

XX FT Synthesis of mature protein or polypeptide - by using bacterial host  
 XX transformed by cloned vehicle contg. DNA fragment etc.

XX PS Example; Fig 3; 34pp; English.

XX CC The closest identifiable promoter for the penicillinase gene in pKT241  
 CC (AAAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was  
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
 CC preproinsulin (see AAAN10033). The closest identifiable promoter for the  
 CC penicillinase gene in pKT218 (AAAN10032) is located in the region 14 to 20  
 CC nucleotides before its translational start signal. In the examples, the  
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
 CC fragment (CB6) for rat preproinsulin (see AAAN10034)

XX SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60  
 |||||  
 Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116  
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
 |||||  
 Db 115 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 75  
 |||||

## RESULT 3

AAAN10031/c  
 ID AAAN10031 standard; DNA; 251 BP.

XX AC AAAN10031;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;  
 XX insulin; ds.

XX OS Escherichia coli.

Key	Location/Qualifiers
FT misc_feature	1..4
FT	/tag= a
FT	/label= sticky end
FT misc_feature	248..251
FT	/tag= b
FT	/label= sticky end

XX EP38182-A.

XX PD 21-OCT-1981.

XX PF 09-APR-1981; 81EP-00301561.

XX PR 11-APR-1980; 80US-00139225.

XX PA (HARD ) HARVARD COLLEGE.

XX PI Gilbert W, Talmadge K;

XX DR WPI; 1981-80125D/44.

XX DR P-PSDB; AAP10038.

XX FT Synthesis of mature protein or polypeptide - by using bacterial host  
 XX transformed by cloned vehicle contg. DNA fragment etc.

XX PS Example; Fig 2; 34pp; English.

XX CC The closest identifiable promoter for the penicillinase gene in pKT241  
 CC (AAAN10031) is located in the region 14 to 20 nucleotides before its  
 CC translational start signal. In the examples, the 3' end of pKT241 was  
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat  
 CC preproinsulin (see AAAN10033). The closest identifiable promoter for the  
 CC penicillinase gene in pKT218 (AAAN10032) is located in the region 14 to 20  
 CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA  
CC fragment (CB6) for rat preproinsulin (see AAN10034)

xx  
SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;  
Query Match 100.0%; Score 101; DB 1; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60  
Db |||||  
175 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 116  
Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 101  
Db |||||  
115 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 75

## RESULT 4

AAV31229/c  
ID AAV31229 standard; DNA; 400 BP.

AC AAV31229;

DT 01-OCT-1998 (first entry)

XX E. coli J96 pathogenicity island contig #43.

XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;  
KW PAI V; phev; vaccine; protective immune response; ds.

XX Escherichia coli.

XX WO9822575-A2.

XX 28-MAY-1998.

XX 21-NOV-1997; 97WO-US021347.

XX 22-NOV-1996; 96US-0031626P.

PR 14-OCT-1997; 97US-0061953P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (UTWI-) UNIV WISCONSIN.

XX Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to develop  
PT products for the detection of pathogenic E. coli and to elicit an immune  
PT response to pathogenic E. coli.

PS Claim 21; Page 140-141; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing  
CC pathogenicity island (PAI) sequences, and represents a nucleic acid  
CC molecule of the invention. PAIs are large fragments of DNA which comprise  
CC pathogenicity determinants. The sequences of the invention are taken from  
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near phev)  
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
CC approximately 94 min (at pHER) on the E. coli chromosome and is  
CC approximately 160 kb in size. Antibodies specific to the proteins encoded  
CC by the PAI open reading frames of the invention can be used in kits to  
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
CC a protective immune response in an animal to the uropathogenic E. coli  
CC strain J96

SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60  
Db |||||  
165 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 106  
Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 101  
Db |||||  
105 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 65

## RESULT 5

AAN60624/c  
ID AAN60624 standard; DNA; 456 BP.

XX AAN60624;

XX 25-MAR-2003 (revised)  
DT 29-OCT-1991 (first entry)

DE Plasmid pUG201 sequence encoding beta-urogastrone.

KW Beta-lactamase signal peptide; pGH54; pGH55; ss.

XX Synthetic.

FH Key Location/Qualifiers

FT promoter 125..170

FT /\*tag= a

FT RBS 200..203

FT /\*tag= b

FT CDS 209..439

FT /\*tag= c

FT sig\_peptide 209..277

FT /\*tag= d

FT /label= Beta-lactamase signal peptide

FT mat\_peptide 278..436

FT /\*tag= e

FT /label= Beta-urogastrone

XX WO8603779-A.

XX 03-JUL-1986.

XX 19-DEC-1985; 85WO-JP000696.

XX 21-DEC-1984; 84JP-00271206.

XX (EART ) EARTH CHEM CO LTD.

PA (OHGA/) OHGAI H.

PI Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

XX WPI; 1986-182911/28.

DR P-ESDB; AAP60678.

XX Recombinant vector for polypeptide secretion - contains signal peptide

PT sequence directly bonded to peptide-coding sequence.

PS Disclosure; Table 4; 79pp; Japanese.

XX The plasmid produces secreted beta-urogastrone in a transformed  
CC expression system. Similar plasmids may be constructed where the  
CC secretion signal may be coupled with eg. somatostatin, insulin, growth  
CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,  
CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to  
CC correct PA field.)

SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;  
Best Local Similarity 100.0%; Pred. No. 2.6e-21;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
Db |||||
173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||
113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73

RESULT 6
AAN71080/C
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
DE
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
Db |||||
173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||
113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73

RESULT 7
AAN71080/C
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
DE
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
Db |||||
173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||
113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/C
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

```

AAN70833/C
ID AAN70833 standard; DNA; 456 BP.
XX
AC AAN70833;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
XX Beta-urogastrone sequence.
DE
XX
XX Tumour; inosine; DNA probe; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= b
FT RBS 200..204
FT /*tag= c
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
XX
XX P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
PT using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
CC polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
CC ssDNA and probe are hybridized and the existence of DNA in the product is
CC detected. It can be used to detect the presence of malignant tumour.
CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
Db |||||
173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db |||||
113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/C
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

```

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 209..277
XX FT /*tag= a
XX FT 278..439
XX FT /*tag= b
XX FT /*product= "New beta-urogastrone deriv."
XX
XX JF63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
XX P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
XX secretion inhibiting action, or cell proliferation promoting action. The
XX deriv. has the same biological or pharmacological activities as beta-
XX urogastrone. It is not susceptible to denaturation by oxidn. and is
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX protease. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGGGTTATTCTCATGAGCGGATACATATTGAATGATTAGAAAAATAACAAATAG 60
DB 173 AGGGTTATTCTCATGAGCGGATACATATTGAATGATTAGAAAAATAACAAATAG 114
QY 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
XX
XX RESULT 9
XX ID ABA90413/c
XX ID ABA90413 standard; DNA; 466 BP.
XX
XX ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antiinflammatory; antipsoariatic; dermatological; antifungal; mitosis;
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX

```







PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241828P.  
PR 01-NOV-2000; 2000US-0244611P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2000US-0254057P.  
PR 05-JAN-2001; 2000US-0259678P.  
XX  
FA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI; 2001-476223/51.  
DR  
XX

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.

Claim 1; SEQ ID NO 418; 512pp; English.

The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders of the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatodystonia, prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic hypertrophy or hyperplasia, and prostate neoplastic disorders, including adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (I) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders, urinary system disorders (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and endocrine disorders (Alzheimer's disease and Parkinson's disease), endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, diseases of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is useful to prevent skin aging, for preventing hair loss, to maintain organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;

Best Local Similarity 100.0%; Pred. No. 2.9e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTTAGAAAATAACAATAG 60  
|||||  
Db 546 AGGGTTATGTCATGAGCGGATACATATTGTAATGTTATTTAGAAAATAACAATAG 487

Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101  
|||||

Db 486 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 446  
|||||

RESULT 15

AAS27819/c

ID AAS27819 standard; DNA; 776 BP.

XX AAS27819;  
XX 07-NOV-2001 (first entry)  
XX DNA encoding novel signal transduction pathway protein, Seq ID 1479.  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
XX anti-inflammatory; anti-HIV; antibacterial; antineoplastic; cancer;  
XX immune system disorder; rheumatoid arthritis; inflammatory condition;  
XX organ transplant rejection; infection; hepatitis C; blood disorder;  
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
XX reproductive system; gastrointestinal; liver disorder; AIDS; ds;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX  
XX WO200154733-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US001312.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225266P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 14-AUG-2000; 2000US-0225759P.  
XX 18-AUG-2000; 2000US-0226273P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226868P.  
XX 23-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0228287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230437P.  
XX 06-SEP-2000; 2000US-0230438P.  
XX 08-SEP-2000; 2000US-0231242P.  
XX 08-SEP-2000; 2000US-0231243P.  
XX 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds  
(without alignments)  
3997.736 Million cell updates/sec

Title: US-09-482-682-8\_COPY\_7369\_7469

Perfect score: 101

Sequence: 1 aggggtattgtctcatgagc.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095 83374 Heb
C 2	101	100.0	300	5	BU963956 EST88 Heb
C 3	101	100.0	300	5	BU964094 EST226 He
C 4	101	100.0	309	9	AL000426 F.rubripe
C 5	101	100.0	391	1	AL597149 DKFZp313J
C 6	101	100.0	414	9	CC819240 100005D19
C 7	101	100.0	417	4	BU684174
C 8	101	100.0	491	9	CC819923 100006J13
C 9	101	100.0	495	4	BI805285 S035A01 S
C 10	101	100.0	495	9	CC818374 100004B07
C 11	101	100.0	496	9	CC818523 100004L13
C 12	101	100.0	503	9	CC819854 100006N08
C 13	101	100.0	515	9	CC817752 100003C16
C 14	101	100.0	518	9	CC817128 100002D21
C 15	101	100.0	519	9	CC817162 100002J19
C 16	101	100.0	519	9	CC817796 100003K14
C 17	101	100.0	521	9	CC819067 100005C09
C 18	101	100.0	533	9	CC819841 100006L07
C 19	101	100.0	542	9	CC816892 100002L01
C 20	101	100.0	550	7	CR766622 DKFZp469H
C 21	101	100.0	551	9	CC816905 100002N02
C 22	101	100.0	554	9	CC819058 100005A09
C 23	101	100.0	563	9	CC819270 100005C21
C 24	101	100.0	566	9	CC816848 100002D02

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954 100002F11
C 29	101	100.0	571	9	CC818423 100004J12
C 30	101	100.0	580	9	CC819098 100005I07
C 31	101	100.0	582	1	AL694813 DKFZp313I
C 32	101	100.0	583	9	CC817633 100003M06
C 33	101	100.0	583	9	CC818436 100004M08
C 34	101	100.0	586	9	CC816883 100002J03
C 35	101	100.0	588	9	CC817788 100003I18
C 36	101	100.0	588	9	CC818340 100004K04
C 37	101	100.0	589	9	CC817595 100003G03
C 38	101	100.0	590	9	CC819754 100006L06
C 39	101	100.0	592	9	CC817679 100003F10
C 40	101	100.0	592	9	CC818508 100004I16
C 41	101	100.0	593	9	CC816942 100002D10
C 42	101	100.0	593	9	CC817699 100003J09
C 43	101	100.0	594	9	CC818287 100004A02
C 44	101	100.0	594	9	CC818422 100004J11
C 45	101	100.0	595	9	CC816929 100002B08

## ALIGNMENTS

RESULT 1  
BM078095/c 300 bp mRNA linear EST 30-NOV-2001  
LOCUS 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma  
DEFINITION  
ACCESSION BM078095  
VERSION BM078095.1 GI:17157967  
KEYWORDS EST.  
SOURCE Hebeloma cylindrosporum  
ORGANISM Hebeloma cylindrosporum  
REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Agaricales; Cortinariaceae; Hebeloma.  
AUTHORS Wipf, D., Benjdia, M., Tegeder, M. and Frommer, W.B.  
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wipf D  
ZMBP - Center for Molecular Biology of Plants  
University of Tuebingen  
Auf der Morgenstelle 1, 72076 Tuebingen, Germany  
Tel: 49 7071 2976160  
Fax: 49 7071 293287  
Email: daniel.wipf@zmbp.uni-tuebingen.de  
PCR Primers  
FORWARD: pDR196 5' primer (PMA 5')  
High quality sequence stop: 300  
POLYA-No.

## FEATURES

Location/Qualifiers  
1..300  
/organism="Hebeloma cylindrosporum"  
/mol\_type="mRNA"  
/strain="H1"  
/db\_xref="taxon:76867"  
/tissue\_type="Mycelia"  
/lab\_host="E. coli XL1-Blue"  
/clone\_lib="Hebeloma cylindrosporum functional cDNA library"  
/note="Vector: pDR 196 (unpublished); Site 1: EcoRI; Site 2: XhoI"

## ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 8.1e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAATAACAATAG 60

```

|||||
174 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c
LOCUS EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10).

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 60
|||||
Db 170 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 111
|||||

RESULT 3
BU964094/c
LOCUS EST226 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU964094
VERSION BU964094.1 GI:24204891
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10).

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 60
|||||
Db 170 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 111
|||||

RESULT 4
FR0009140
LOCUS F.ru.ribripes GSS sequence, clone 010H20aC4, genomic survey sequence.
DEFINITION F.ru.ribripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.

FEATURES
Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 60
|||||
Db 170 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 111
|||||

RESULT 4
FR0009140
LOCUS F.ru.ribripes GSS sequence, clone 010H20aC4, genomic survey sequence.
DEFINITION F.ru.ribripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 98
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 139
RESULT 5
AL597149 391 bp mRNA linear EST 04-SEP-2003
LOCUS DKFP313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DEFINITION DKFP313J1611 5', mRNA sequence.
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.
JOURNAL EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
COMMENT Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMF2 (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFP313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFP313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: SfIIA; Site_2: SfIIB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 60
Db 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 355
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 98
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 139
RESULT 6
CC819240/c 414 bp DNA linear GSS 17-JUL-2003
LOCUS CC819240
DEFINITION CC819240.1 GI:32899308
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION CC819240
VERSION CC819240.1
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 414)
AUTHORS Dunn, D., Doak, T., Herrick, G. and Weiss, R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL Unpublished (2003)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 414.
Location/Qualifiers
    1..414
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC10005D19"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /note="Vector: PWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adaptor mouse DNA was annealed to
        adaptor vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAACAATAG 355
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 354 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 314

```

```

RESULT 7
BJ684174/c
LOCUS BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
RNA sequence.
ACCESSION BJ684174.1 GI:46527295
VERSION BJ684174
KEYWORDS EST.
SOURCE Haplochromis chilotes
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
AUTHORS Watanabe,M., Kobayashi,N., Shin-I,T., Kohara,Y. and Okada,N.
TITLE Orf sequences of cichlid in Lake Victoria are essentially same
JOURNAL Unpublished (2004)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehinigenes.nig.ac.jp.
FEATURES
Location/Qualifiers
1..417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stage="varied"
/clone_lib="HCEST library"

ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
Db 129 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 70

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 101
Db 69 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 29

RESULT 8
CC819923/c
LOCUS CC819923 491 bp DNA linear GSS 17-JUL-2003
DEFINITION 100006J13R Oxytricha plasmid UUGC100006J13 R, genomic survey
histriomuscorum genomic clone UUGC100006J13 R, genomic survey
sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL Unpublished (2003)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: ddm@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGGAACACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES
Location/Qualifiers
1..491
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100006J13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
Db 412 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 353

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 101
Db 352 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 312

RESULT 9
BI805285
LOCUS BI805285 495 bp mRNA linear EST 02-OCT-2001
DEFINITION S035A01 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 495)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES
Location/Qualifiers
1..495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

```



/tissue type="Stem"  
/dev stages="3-5 leaf stage"  
/clone lib="Stem library from Oryza sativa (3-5 leaf stage)"  
/note="Vector: pSport2"

## ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 495;  
Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 60  
|||||  
Db 62 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 121  
|||||

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGTC 101  
|||||  
Db 122 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGTC 162  
|||||

## RESULT 10

CC818374/c  
LOCUS CC818374.1 GI:32897661 495 bp DNA linear GSS 17-JUL-2003  
DEFINITION 100004807R Oxytricha plasmid UUGC100004B07 R, genomic survey  
histriomuscorum sequence.

## ACCESSION

CC818374

## VERSION

CC818374.1

## KEYWORDS

GSS

## SOURCE

Sterkiella histriomuscorum (Oxytricha trifallax)

## ORGANISM

Sterkiella histriomuscorum

Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;

Stichotrichida; Oxytrichidae; Sterkiella.

1 (bases 1 to 495)

Dunn, D., Doak, T., Herrick, G. and Weiss, R.

Paired end reads from plasmid inserts of Oxytricha trifallax

macronuclear chromosomes

Unpublished (2003)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Plate: 0004 row: B column: 07

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 495.

Location/Qualifiers

1..495

/organism="Sterkiella histriomuscorum"

/mol\_type="genomic DNA"

/db\_xref="taxon:94289"

/clone="UUGC100004B07"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone\_lib="Oxytricha plasmid UUGC10 library"

/note="Vector: PWD42nv; Purified macronuclear chromosomal

DNA from Oxytricha trifallax was blunt end-repaired with

T4 DNA polymerase and T4 polynucleotide kinase. Adaptor

oligonucleotides were ligated to the blunt ends in high

molar excess. Vector DNA was prepared from a derivative of

PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible

derivative of plasmid R1. The vector was ligated with

adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. Coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

## ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 495;

Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 60  
|||||

Db 392 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 333  
|||||

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGTC 101  
|||||

Db 332 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGTC 292  
|||||

## RESULT 11

CC818523/c

LOCUS CC818523.1 GI:32897943 496 bp DNA linear GSS 17-JUL-2003

DEFINITION 100004L13R Oxytricha plasmid UUGC100004L13 R, genomic survey

histriomuscorum sequence.

CC818523

CC818523.1

GI:32897943

GSS

SOURCE

Sterkiella histriomuscorum (Oxytricha trifallax)

Sterkiella histriomuscorum

Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;

Stichotrichida; Oxytrichidae; Sterkiella.

1 (bases 1 to 496)

Dunn, D., Doak, T., Herrick, G. and Weiss, R.

Paired end reads from plasmid inserts of Oxytricha trifallax

macronuclear chromosomes

Unpublished (2003)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Plate: 0004 row: L column: 13

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 496.

Location/Qualifiers

1..496

/organism="Sterkiella histriomuscorum"

/mol\_type="genomic DNA"

/db\_xref="taxon:94289"

/clone="UUGC100004L13"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone\_lib="Oxytricha plasmid UUGC10 library"

/note="Vector: PWD42nv; Purified macronuclear chromosomal

DNA from Oxytricha trifallax was blunt end-repaired with

T4 DNA polymerase and T4 polynucleotide kinase. Adaptor

oligonucleotides were ligated to the blunt ends in high

molar excess. Vector DNA was prepared from a derivative of

PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible

derivative of plasmid R1. The vector was ligated with

adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. Coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

## ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 496;

Best Local Similarity 100.0%; Pred. No. 8.4e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 60  
|||||

Db 391 AGGGTTATTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAATAAACAATAG 332  
|||||

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGTC 101  
|||||

## RESULT 13

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Stichotrichida; Oxytrichidae; Sterkiella.  
 1 (bases 1 to 518)  
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.  
 Paired end reads from plasmid inserts of Oxytricha trifallax  
 macronuclear chromosomes  
 Unpublished (2003)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Plate: 0002 row: D column: 21  
 Seq primer: CACACAGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 518.

FEATURES  
 source  
 1..518  
 Location/Qualifiers  
 /organism="Sterkiella histriomuscorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:94289"  
 /clone="UUGC100002J19"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Oxytricha plasmid UUGC10 library"  
 /note="Vector: PWD42nv; Purified macronuclear chromosomal  
 DNA from Oxytricha trifallax was blunt end-repaired with  
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
 oligonucleotides were ligated to the blunt ends in high  
 molar excess. Vector DNA was prepared from a derivative of  
 pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
 derivative of plasmid R1. The vector was ligated with  
 adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. Coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60  
 Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 351  
 Qy 61 GGGTTCCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101  
 Db 350 GGGTTCCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 310

RESULT 15  
 CC817162/c  
 LOCUS  
 DEFINITION  
 histriomuscorum genomic clone UUGC100002J19 R, genomic survey  
 sequence.  
 CC817162  
 CC817162.1 GI:32896449  
 GSS  
 Sterkiella histriomuscorum (Oxytricha trifallax)  
 Sterkiella histriomuscorum  
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;  
 Stichotrichida; Oxytrichidae; Sterkiella.  
 1 (bases 1 to 519)  
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.  
 Paired end reads from plasmid inserts of Oxytricha trifallax  
 macronuclear chromosomes  
 Unpublished (2003)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center

University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Plate: 0002 row: J column: 19  
 Seq primer: CACACAGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 519.

FEATURES  
 source  
 1..519  
 Location/Qualifiers  
 /organism="Sterkiella histriomuscorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:94289"  
 /clone="UUGC100002J19"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Oxytricha plasmid UUGC10 library"  
 /note="Vector: PWD42nv; Purified macronuclear chromosomal  
 DNA from Oxytricha trifallax was blunt end-repaired with  
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor  
 oligonucleotides were ligated to the blunt ends in high  
 molar excess. Vector DNA was prepared from a derivative of  
 pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible  
 derivative of plasmid R1. The vector was ligated with  
 adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. Coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

ORIGIN  
 Query Match 100.0%; Score 101; DB 9; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60  
 Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 357  
 Qy 61 GGGTTCCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101  
 Db 356 GGGTTCCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:22:32  
 Job time : 963.667 secs

**THIS PAGE BLANK (USPTO)**